Examiner feedback and learning: What are the characteristics of effective remote feedback in a hierarchic, professional context?

Martin Joseph Johnson

Corpus Christi College

June 2017

This dissertation is submitted for the degree of Doctor of Philosophy

Faculty of Education – University of Cambridge

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

It does not exceed the prescribed word limit of 80,000 words.

Martin Joseph Johnson

Acknowledgements

I will always be thankful to my supervisor, Professor Neil Mercer, for his guidance, expertise and inspiration. Through Neil's careful support I have developed my thinking in directions that I never envisaged when I started this study. It is fair to say that I left every supervision feeling better than when I entered it. I also want to thank my advisor, Paul Warwick, for the way that he has attended closely to the fine details of my thinking, as well as making me laugh. Together, the feedback from Neil and Paul has encouraged me to grow as an academic thinker.

I want to say thanks to Gerry, who has always, and from the very start, encouraged me to get on with more learning, and to Oli and Stan, who always remind me of the important things in life and help to keep my perspective in balance.

I would like to mention all the inspirational people that I have met whilst undertaking this study, from coders in Cyprus to criminologists in Copenhagen; it's been an interesting time. In particular, I need to thank the examiners who were a central part of this research and who I enjoyed spending time with and learning from. I would also like to thank Beth Black for all the conversations that we had and which helped me identify this as an area of study interest. Thanks also to Tori Coleman for helping me to knock my thesis into some shape, and to Karen Barden for helping me through the particularities of the University finance system.

And finally, I want to thank Cambridge Assessment and Corpus Christi College for the funds that have supported my study. I very much appreciate the contribution that you have made to my personal development.

Table of Contents

List of Tables and Figures

1. Introduction

1.1 Aims

My study explores the characteristics of remote feedback that examiners communicate to each other; with the aim of isolating those features that effectively help examiners to align their thinking. The problem at the core of the research question is shaped by the context in which the study is situated. My study focuses on examiners' practices in the Oxford, Cambridge and RSA (OCR)[1] Awarding Body. Like other large scale educational Awarding Bodies, the principal ambition of OCR is to administer examinations that deliver fair and equitable outcomes. To this end, the reliability of an examiner's marking is a key indicator that an assessment system is achieving such ambitions. For an examiner to mark reliably they need to be able to interpret and apply a mark scheme consistently with the standard established by the most senior examiner in their marking panel. To support this ambition a number of quality assurance arrangements exist, including mechanisms for senior examiners (known as Team Leaders [TLs]) to give feedback to examiners in their team.

I argue that examiner feedback interactions represent a potentially powerful learning opportunity. In accordance with quality assurance arrangements, examiner interaction takes place in a hierarchically structured context, with more expert TLs being responsible for monitoring the marking quality of a team of more junior examiners. As part of this monitoring role, TLs are expected to give regular feedback to the examiners in their team. Through these interactions with more expert senior colleagues, examiners gain insights into the marking procedures and linguistic interpretations that characterise a professional examiner community of practice. In this context, learning involves the examiner becoming adept at using particular tools and privileging certain forms of thinking in applied situations. This conceptualisation of learning has been characterised by Lave & Wenger (1991) as the movement from the periphery towards the centre of a professional community and involves a process of cognitive apprenticeship (Dennen, 2004) that builds on Vygotskian concepts of thought and development (e.g. Vygotsky, 1978). As a consequence, I adopt a sociocultural model of learning for this study that suggests that education is 'an interpersonal and

[1] I have tried to explain the meaning of particular assessment industry-related or technical terminology when it occurs through the thesis. I have also included a Glossary of Terms in Appendix A.

intrapersonal process' (Mercer, Littleton, & Wegerif, 2004, p.203) mediated by cultural tools (e.g. computer-based messaging and specialised language).

In an effort to reinforce and support the ambition to deliver fair and equitable marking outcomes, this research identifies and evaluates the characteristics of feedback that help to align examiners' marking. In so doing, the study gives insight into the complexity of feedback giving and contributes to the literature on expertise in this field, particularly in contexts where there is asymmetry in expertise and relative professional status. This has implications that stretch beyond the immediate study context, with relevance for anyone who is responsible for delivering performance feedback in professional organisations.

The study also provides insight into the potential hurdles to learning in hierarchic contexts, where access to powerful knowledge is asymmetrically distributed. It is recognised elsewhere that the forms of communication that are encouraged within institutions, including feedback communication, influence the meaning-making processes that take place within them, since they 'serve up a limited register of typifications (words) that can be used to construct a course of action' or notice certain problems (Weber & Glynn, 2006, p. 1649). This coheres with observations from other, safety-critical and high-risk professional environments, such as police firearms training or healthcare systems, where some learning approaches have been previously conceptualised as being prone to inertia and conservatism (Beighton, Poma, & Leonard, 2014; Wiredu, 2007).

1.2 Background to the Study: The Research Context

1.2.1 Team Leader [TL] Feedback Requirements

In England, Wales and Northern Ireland, educational qualifications are offered by designated Awarding Bodies that have been recognised as being eligible to award such qualifications by the national body that regulates qualifications and examinations (the Office of Qualifications and Examinations Regulation [Ofqual]). In order to satisfy the regulator's eligibility standards, an Awarding Body needs to act in accordance with the regulator's Code of Practice. This Code sets out the conditions which should inform the Awarding Body's quality assurance (QA) processes to ensure that they 'promote quality, consistency, accuracy and fairness' through their assessment and awarding processes (Ofqual, 2011, p.5). As such, the examiners who are the focus of this study work in a highly regulated environment.

A stipulation of the Ofqual Code of Practice is that Awarding Body QA arrangements are based on hierarchic principles. As a consequence, the responsibility for developing mark schemes and monitoring examiners' marking quality is invested in those at the apex of the hierarchy. This hierarchy is structured around key assessment roles, such as Principal Examiners and Team Leaders (TLs). At the base of the system are Assistant Examiners (otherwise known simply as 'examiners').

The hierarchic nature of the QA process is reinforced in the Code's expectation that:

> *'…all examiners have a well-founded and common understanding of the requirements of the mark scheme (appropriate to their responsibilities) and can apply them reliably. This is the responsibility of the principal examiner, whose professional judgements on the interpretation and application of the mark scheme for the unit/component must be final.' (Ofqual, 2011, p.25)*

This last statement is highly significant as it articulates the regulator's conceptualisation of high quality marking. Marking quality is measured through reliability, with an implicit assumption that all examiners must mark as though they are the Principal Examiner. As a result, observed marking discrepancies between examiners assume that any error lies with least senior examiners.

To conform to regulatory requirements the OCR Awarding Body operates QA arrangements in two phases, with feedback having an important role (Figure 1.1). These QA arrangements are facilitated by a digital marking environment (*scoris® assessor*). This digital marking environment hosts students' examination scripts as digital objects and enables these script images, and any attached marking data, to be distributed efficiently between examiners and TLs.

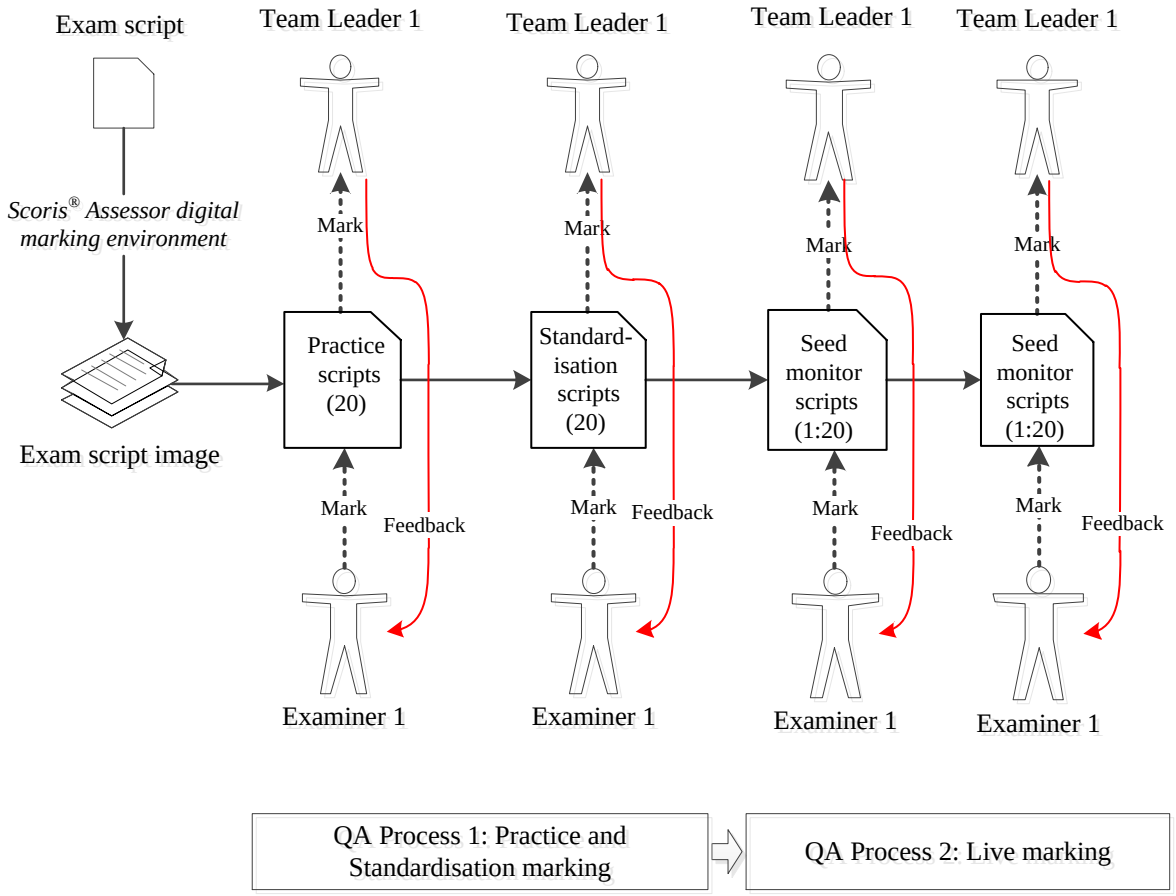

Figure 1.1: The OCR Marking QA Process

In the first QA phase each examiner marks a group of common practice scripts. These are a sample of student exam scripts that have been chosen by the senior examiner team because they exemplify particular elements of the mark scheme. Following this practice marking, each examiner receives feedback from their TL on their marking performance and then marks a group of common standardisation scripts. Again, this script sample has been chosen by the senior examiner team because they exemplify performances that are in some way representative of the scripts that examiners will mark in the operational, or live, marking phase. If the examiner's marking on these standardisation scripts is in close agreement with the predefined marks given to them by the senior examiner team, the examiner is deemed qualified to mark live exam scripts.

The second QA phase takes place during live marking. One in every 20 scripts that an examiner receives to mark has been pre-marked by a group of senior examiners. These scripts are known as seed scripts. The examiner is unaware of which scripts are pre-marked, but these seed scripts allow the TL to monitor the alignment levels of an examiner mark with a definitive pre-mark. TLs are expected to give feedback to the examiners on their seed script marking, either through electronic mail (through the *scoris® assessor* system or through personal electronic mail accounts) or via the telephone.

1.2.2 Team Leader [TL] Feedback Practices

Having outlined the formal processes pertaining to the examination marking process in OCR, it is important to consider what is already known about how examiners actually work within these requirements. Previous work that I have been involved in used observation methods to capture senior examiner feedback practices. These studies highlight some of the complexity of TLs' examiner monitoring and feedback work, including that they commonly pass through evaluation and diagnosis stages prior to giving feedback to examiners (Johnson, 2015; Johnson & Black, 2012b). These stages allow the TL to consider the potential reasons for an examiner's mark and to make decisions about how best to give feedback to remediate any marking issues. The complexity of TL feedback giving also involves a number of other factors. Some TLs consider the potential affective impact of feedback messages on the recipient (Johnson & Black, 2012a), with feedback communication mode choices being influenced by the character of the communication being conveyed (Johnson, 2016a).

A final complexity also pertains to the affordances of the communication arrangements in the OCR professional examiner context. Johnson & Black (2012b) report how the *scoris® assessor* digital marking environment makes it easy for examiners to respond directly to TL feedback with the click of a button. In addition, examiners can instigate communication by contacting their TL through the system's electronic messaging function. This iterative facility highlights the two-way potential of feedback communication in the *scoris® assessor* environment, which may be in tension with the hierarchic assumptions of the QA arrangements. This has been demonstrated in my earlier work, which showed TLs displaying equivocation or altering their perspective on a marking decision through discussion with less senior examiners (Johnson & Black, 2012b).

This tension between a feedback model that transmits and reinforces a definitive mark down the examiner hierarchy and the evidence that mark meaning may be sometimes negotiated

within feedback communication may have an ontological root. Johnston (2004) argues that the concept of a true mark that captures the qualities of a performance is underpinned by a positivist epistemological stance. In this conceptualisation 'expert examiners can use their expertise to make a judgement which is assumed to come close to an objective reality' (Johnston, 2004, p.398). This perspective contrasts with an interpretivist position, which suggests that a social reality (such as a true mark) is a social construction rather than an absolute objective truth. Balancing these positions, it would be anticipated that at least some marking judgements could involve examiners sharing their interpretative perspectives as part of a negotiated consensual social process.

Again, there is evidence from earlier research that suggests that disagreement and interpretation is a feature of marking. Moreover, this research also suggests that disagreement relates particularly to the level of complexity of some types of items (exam questions). It is possible to define items by their characteristics, (for example, see the framework of Black, Curcin, & Dhawan, 2010[2]). Massey & Raikes (2006) found item type complexity to be an influence on examiner agreement levels, with more objective items being subject to more agreement than other items. Bramley (2008) goes on to argue that this disagreement relates to the potential level of constraint that pertains to the item, with less agreement being found where there is less constraint on the acceptable answers allowed by the item.

A final influence on examiner agreement is the nature of the mark scheme and the space for any potential ambiguity that it possesses. The semiotic qualities of text make multiple readings possible, and this holds true of mark schemes even though they undergo various quality controls. As Wiliam (1993) argues, 'no criterion, no matter how precisely phrased, admits of an unambiguous interpretation' (Wiliam, 1993, p341).

Taken together, these studies are a spur for my current research, as they hint at some of the factors that contribute to the complexity of TL feedback practices. These complexities suggest that TLs not only consider the 'what' of feedback content, but also the 'how' of feedback delivery. My initial work led me to want to explore further the nature of the

[2] ***Objective items***: Candidates are required to provide a highly defined response such as adding a line to a diagram, circling the correct option etc. No variation in responses is allowed by the mark scheme. ***Constrained items***: Candidate needs to generate their own specific word or words. Little or no variation in response is cued. [e.g. 'State']. ***Short answer questions***: A form of free response where the required answer should be succinct, e.g. a phrase or several sentences (and where the mark scheme allows some variation in responses). [e.g. 'Explain', 'Suggest']. ***Extended responses***: Candidates have to create a given amount of text and/or working; candidates need to organise their responses.

feedback that was communicated to examiners, as well as what might be considered effective about this content.

This exploration entails developing a theoretical framework for broadly conceptualising the feedback process. Such a framework has to take into consideration as many elements of the TL feedback task and environment as possible, in order to anticipate influences on feedback discourse and to therefore inform a data gathering strategy for the research study. The next section of this chapter outlines the structure of my thesis, before presenting an overview of the research design.

1.3 Structure of the Thesis

I structure my thesis into seven chapters, including this introductory chapter, which outlines core ideas and intentions. My research explores two interlinked research questions: 'What are the characteristics of examiner feedback?', and 'What are the characteristics of effective examiner feedback?' In order to address the second question I hypothesise that effectiveness relates to how feedback features support or undermine examiners' common ground building.

Chapter 2: In this chapter I present a literature review that brings together feedback research that is located in applied, remote and hierarchic contexts. The aim of this review is to isolate any feedback factors that are considered to be effective. For the purposes of this study, I consider effective feedback to be communication that helps participants to align their thinking and action. This review culminates in my development of an Integrated Analytical Framework (IAF) that includes 10 dimensions. These dimensions are used in a later chapter (Chapter 6) to theoretically ground my evaluation of the gathered feedback data to address the research question: 'What are the characteristics of effective examiner feedback?'

Chapter 3: In this chapter I outline the context of the study as a precursor to explaining the theoretical perspective that I develop for my study. Drawing on sociocultural learning theories I suggest that feedback is a collaborative process where participants develop and maintain common ground, and that it is not an unproblematic transmission of information. In constructing my framework I also look beyond conventional sociocultural theories to consider how discourse is organised, and use the concept of 'articulation' (Strauss, 1985) which describes how communication helps to coordinate interconnected work across individuals.

Finally, I suggest that the articulation work of building and maintaining common ground allows the examiners to create an Intermental Development Zone (IDZ) where their minds are mutually attuned.

Chapter 4: In this chapter I give a detailed outline of the methodological framework that I adopt for my study. I start with a description of my pilot study and how this influences the methods that I adopt for the main study. I then go on to describe the methods that I use in my main study, structuring this discussion around two research questions: 'What are the characteristics of examiner feedback?', and 'What are the characteristics of effective examiner feedback?'

For the first research question I develop a methodology that extends the Sociocultural Discourse Analysis (SCDA) approach developed by Mercer (2004) and Littleton & Mercer (2013); I call this approach 'augmented Sociocultural Discourse Analysis' (ASCDA). My methodology allows me to investigate the features of interaction at both a particular and a general level, and clusters my analysis into four specific feedback discourse themes: feedback content, the development of discourse over time, evidence of *Joint Intellectual Action* within feedback, and the impact of feedback. Within this chapter I justify how I construct my ASCDA methodology by integrating a group of methods that are generally associated with discrete approaches to discourse analysis (Thematic Content Analysis, Conversation Analysis, and Corpus Linguistics).

For the second research question, I identify dimensions that influence feedback effectiveness. This process involves case study interviews with a sample of examiners. The focus of this part of the study is to engage the participants in self-reflection about their feedback experiences around a group of common messages. For these case studies I gather interview evidence using a Stimulated Recall technique. This process allows me to consider both the TLs' intentions underpinning the delivery of specific feedback messages and examiners' reactions to those messages.

Chapter 5: In this chapter I present my findings for the first research question: 'What are the characteristics of examiner feedback?' In the first section of this chapter I outline the characteristics of the 991 instances of examiner feedback data that I gathered over two marking sessions. I then go on to analyse these data using the ASCDA approach described in Chapter 4 and report my findings according to the four themes that I associate with the methodology.

Chapter 6: In this chapter I present my findings for the second research question: 'What is effective feedback?' My approach to the analysis of effectiveness has two parts, one that is theoretically driven and another that is problem-oriented and empirically driven. For the first part of my analysis I relate the characteristics of feedback that I report in Chapter 5 to the research literature that I organised into an Integrated Analytical Framework (IAF) (outlined in Chapter 2). For this analysis I draw on the interview data from a sample of participants who self-reflect on their intentions behind, and reactions to, particular feedback messages.

The second section of my analysis is oriented to the problem of how participants actually attain convergence through feedback. This section is organised around a series of cases that exemplify how the research participants engage in extended episodes of feedback to reach convergence. In this way the feedback centres on a shared problem (an examination question) and the analysis illustrates how divergence is diminished through the participants' actions.

Chapter 7: In this final chapter I synthesise the findings from Chapters 5 and 6 to consider the lessons for examiner practice in particular, and for other professional feedback practices more generally. Taken together, these analyses suggest that TLs fulfil an intellectually challenging process when giving feedback to examiners. My analyses suggest that this complexity involves the TLs establishing and maintaining an Intermental Development Zone (IDZ) with examiners through their feedback communication, and that this entails them manipulating discourse features whilst simultaneously attending to a variety of contextual features of the professional environment.

In the next section I give a brief overview of my research design.

1.4 Research Design Overview

Following a pilot phase, I gathered the feedback messages from three TLs and their teams over two marking sessions (2014 and 2015). Three academic disciplines were chosen as the focus for this study because I wanted to increase the external validity of the research design, expanding the generalizability of any findings beyond any specific subject domain or particular marking team community. The chosen sample of TLs included a mix of those

marking UK Advanced level General Certificate of Education (GCE) subjects[3] (Chemistry, Economics and Geography[4]). These subjects were chosen because they included scripts that incorporated some subjective items. These items tend to invite performances that require higher order skills, and inevitably involve complex decision making on the part of an examiner when applying the mark scheme. Such items are considered to be the most complex item type, and I anticipated that they would produce rich between-examiner interactions.

The feedback messages that I gathered during the study are a mixture of electronic mail (*scoris® assessor* mail or personal email) and transcribed telephone messages. Figure 1.2 is a visual representation of the main study design.

The feedback and interview data that I collected in Phase 1 are augmented by additional data gathering in Phase 2, to increase confidence to my analysis. The feedback data gathered in each phase was analysed according to the ASCDA approach that I developed specifically to generate a description of examiner feedback. This strand of analysis and how it relates to my data collection approach is shown by the black lines in Figure 1.2. My description of the gathered data is then evaluated against a framework that explores those elements considered to be effective in terms of supporting the alignment of TLs' and examiners' perspectives. This evaluation, shown by the dashed lines in the Figure, involved the consideration of participants' self-reflection interview data that I considered in relation to an analytic framework (IAF) based on my literature review.

[3] GCE courses are usually studied over a two-year period and are widely recognised in England, Wales and Northern Ireland as being the standard entry qualification for assessing the suitability of applicants for academic courses in UK Universities. GCE students are usually around 18 years old.

[4] ***Chemistry A [Unit Code F322]***: Chains, Energy and Resources. 100 mark paper; ***Economics [Unit Code F581]***: Markets in Action. 60 mark paper; ***Geography [Unit Code F581]***: Managing Change in Human Environments. 75 mark paper.

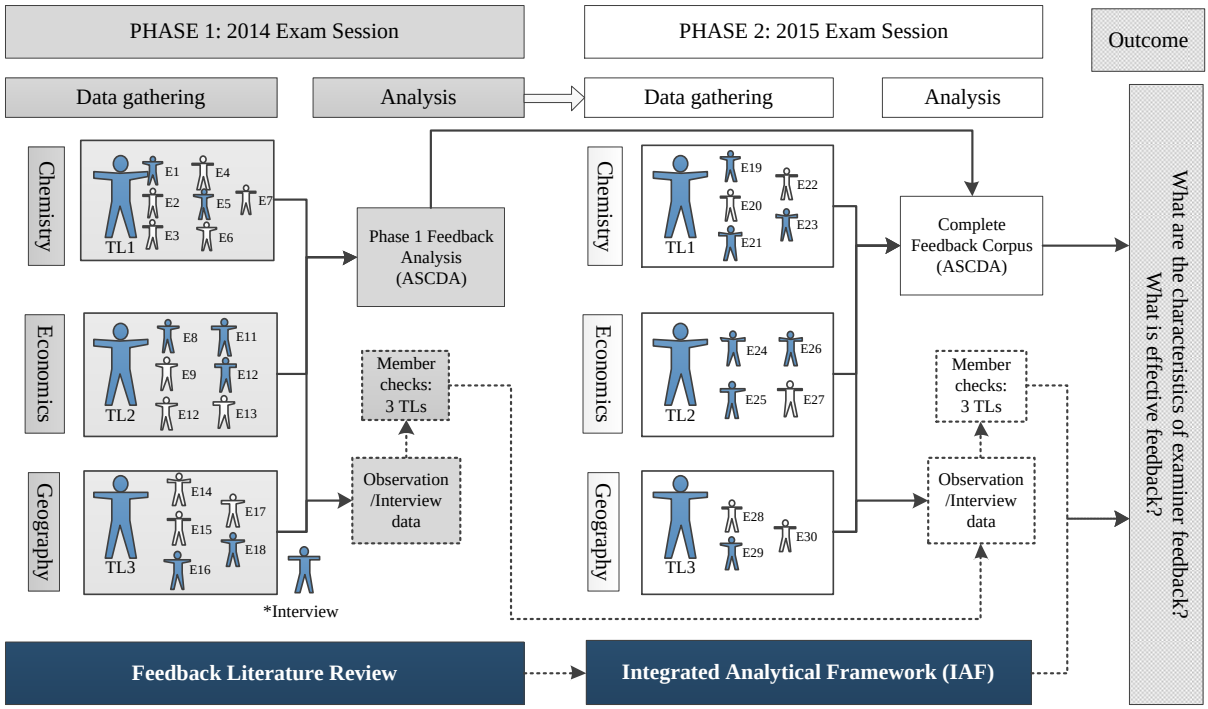

Figure 1.2: Project Design showing the Phases of the Research Process

The next chapter provides a literature review of research on the nature and role of feedback in learning.

2. Literature Review

2.1 What Factors Influence Feedback Effectiveness?

There is an extensive educational research literature that incorporates reflection on the role of feedback in learning (as is evidenced in the review outcomes presented in Table 2.3 below). This literature attests to the idea that feedback is of significant interest for those who seek to understand learning processes. This also appears to be the case for professional learning, with feedback being a form of institutional discourse that influences the types of practices that come to be considered as normal within an organisation (Asmuß, 2008; Kupferberg, 2013). In this way, feedback is an influence on professional practice (Barton & Wolery, 2007).

A relatively common sentiment found in the literature is that feedback training is possible and should have a positive impact on both givers and receivers (Bearman, Molloy, Ajjawi, & Keating, 2013; Guichon, Bétrancourt, & Prié, 2012; Hudson, 2014; Hyland, 2001; Nestel, Bello, & Kneebone, 2013). Despite this, literature highlights that feedback is complex and often under theorised (Archer, 2010; Evans & Butler, 1992), with little research focusing on how to provide feedback (Ahmed et al., 2012; Li & De Luca, 2014; Thurlings, Vermeulen, Kreijns, Bastiaens, & Stijnen, 2012). For Tuck (2012), this oversight is significant since it overlooks the complexity of feedback giving as a social practice.

There is debate about whether feedback constitutes a discourse genre. Building on the work of Swales (2014), Bhatia (1993) considers a genre to be ‘a recognisable communicative event characterised by a set of communicative purposes mutually understood by members of the community in which it occurs’ (Bhatia, 1993, p. 13). There is some general agreement that feedback has the characteristics of a genre (Mirador, 2000; Yelland, 2011), with Mathisen (2012) suggesting that a key to understanding feedback is the ability of a recipient to master the genre (Mathisen, 2012).

There have been several attempts at generating a typology of feedback. These include Pryor & Crossouard's (2008) conceptualisation of convergent and divergent feedback, and Hattie & Timperley's (2007) categorisation of four levels of feedback (task, processing, self-regulation, and person). Whilst both of these influential frameworks have significance, with both giving language to a concept that is ubiquitous in learning discourse, they are also very broad in their scope. For the purposes of my study I wanted to explore feedback effectiveness from a

more particular perspective, looking specifically at the feedback that takes place in professional, hierarchic, and remote contexts.

This is an area of particular interest for me because it describes the conditions in which examiners communicate feedback to each other in the professional context of my study. It is also of interest because some commentators suggest that remote and hierarchic communication can help participants to establish a strong shared identity (Wilson & Williams, 1977). At the same time, it is acknowledged that remote and hierarchic communication can also interfere with the connection between feedback message intention and its reception, due to power-based differentials (e.g. see Lingard, Schryer, Garwood, & Spafford, 2003) and technical considerations (e.g. see Walvoord, Redden, Elliott, & Coovert, 2008). Moore (2007), Murphy & Rodriguez (2008), and Vonderwell (2003) suggest that the concept of transactional distance can be used to describe the psychological and communicative space of potential misunderstandings between teachers and learners. It is also argued that the transactional distance can increase where a teacher controls the flow of feedback information, or that it can be reduced where dialogue exists in feedback (Rovai, 2000).

To explore the literature I carried out a semi-systematic review of feedback research in hierarchic work-based and/or remote contexts. I then organised the messages gleaned from the review into a framework of 10 factors that the literature suggested could influence feedback effectiveness. I call this framework an Integrated Analytical Framework (IAF). I then used the IAF factors as a tool for analysing my own gathered feedback data (See Chapter 6).

2.2 Semi-Systematic Literature Review Method

Evidence suggests that there are a 'limited number of studies on the actual practices of feedback' (Li & De Luca, 2014, p.13), with a 'paucity of empirical evidence of what type of feedback works' (Evans, 2013, p. 73). One approach to supplementing this area is the use of a systematic literature review. I adopted this type of method because it provides a transparent structure for gathering and analysing literature, therefore supporting the potential replicability of the review outcomes. Evans & Benefield (2001) state that the key features of a systematic review include the formulation of an explicit research question; transparency of method;

exhaustive searches that seek both published and unpublished studies; clear criteria to assess the quality of studies; clear criteria for excluding studies based on the scope of the review; joint reviewing to reduce potential bias; and a clear statement of the findings.

I employed most of these review features to gather literature that looks specifically at how feedback is used in hierarchically structured, professional work contexts. I used three different search types: a journal content search, a research database search and an author search. I then augmented this with a snowball search strategy. This is a technique that searches for references of references, and is 'especially powerful for identifying high quality sources in obscure locations' (Greenhalgh & Peacock, 2005, p. 1065), thereby helping to fill the information gaps in obscure social situations through 'chain referral' (Atkinson & Flint, 2001). This technique has been used in other similar studies (e.g. May, 2013; Black & Wiliam, 1998).

I was selective in the elements of Evans & Benefield's (2001) systematic review method that I adopted for a number of reasons. I chose not to adopt an explicit strategy to evaluate the weight of the findings for each study in the review since this aspect of the method reflects its roots in evaluations of efficacy in medical intervention studies. I was concerned that such a focus on intervention and efficacy might overlook theoretically well-grounded studies that involve small numbers and which struggle to demonstrate significant outcomes.

I used Ramaprasad's (1983) widely cited definition of feedback as a basis for the review. Ramaprasad defines feedback as 'information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way' (Ramaprasad, 1983, p.4). I excluded articles that failed to address this notion of feedback. I also established temporal boundaries for literature inclusion. There is a heritage of feedback-related literature that pre-dates interests in technologically-mediated feedback, therefore I chose to include literature that post-dates Sadler's (1989) work on identifying the constituents of formative feedback, as this is a seminal work.

A key feature of a systematic review is the inclusion of an explicit research question (Evans & Benefield, 2001). I was interested in identifying those feedback characteristics that are effective in aligning perspectives in hierarchic, remote interactions, therefore, my analysis involved collating the main research outcomes from across a variety of studies to understand the impact of individual feedback features (e.g. 'what does the research literature say about

how feedback timing influences feedback effectiveness?'). The first stage of this process involved setting a number of search parameters (Table 2.1).

Table 2.1: Literature Search Parameters

Title	Keyword
Feedback +	Work-based Hierarch* Examiner Remote/Distan* Align*

The first four search terms focus on the professional context of the feedback interactions, whilst the final term focuses on the learning processes involved. Table 2.2 outlines the literature search process.

Table 2.2: The Literature Review Process

Stage 1 - Targeted Journal Searches	53 journals based on literature gathered across various academic domains (education, technology, communication, organisational research, assessment) (Appendix 1).
a. Search each journal for feedback [title]. Review of abstract for inclusion. b. Secondary advanced search for feedback [within article] with a link to keywords. Relevant articles were included in the review.	
Stage 2 - Targeted Database Searches	10 Publisher databases and 3 educational research databases across a diverse geographical spread.
a. Search of identified databases for feedback [title]. b. Secondary advanced search for feedback [title] with a link to keywords. Relevant articles were included in the review. c. Any journals that appeared to be particularly pertinent were added to a second round of the Targeted Journal Review process (Stage 1).	
Stage 3 - Targeted Author Search	Using previous literature review outcomes to identify 26 authors who have a multiple presence (i.e. appearing more than twice in the earlier review stages).
a. Internet search of personal/institutional webpages, LinkedIn, Google Scholar, and Academia.edu websites to identify additional articles and conference papers with feedback [title]. Relevant articles were included in the review. b. Any journals that appeared to be particularly pertinent were added to a third round of the Targeted Journal Review process (Stage 1).	
Stage 4 – Snowball search	Additional search based on highly cited articles that have particular relevance to multiple keywords or that also cite articles that have particular relevance to multiple keywords.
a. References cited in relevant journal papers from Stages 1-3 were identified and included.	

I located 183 articles through the four search strategies. Table 2.3 outlines the profile of the articles that were located according to each search term domain in each search stage. The data presented in the table shows that most of the located articles related to remote/distance feedback or work-based feedback.

Table 2.3: Literature Search Results

Stage	Search	[title]	[keyword]					TOTAL
		Feedback	Work-based	Hierarch*	Remote/ Distan*	Examiner	Align*	
1	Journal	496	34	10	51	1	7	103
2	Database	191062	8	6	15	6	2	37
3	Author	90	6	2	5	0	1	14
4	Snowball	-	8	3	6	2	10	29
	TOTAL	191648	56	21	77	9	20	183

Table 2.3 also shows that most of the articles were identified through the journal search (103 articles). The additional database, author, and snowball searches resulted in the identification of an additional 80 articles. These additional searches increased the representation of articles that considered feedback more broadly beyond remote/distance feedback and work-based feedback, and therefore validated my use of a mixture of search strategies.

2.3 Analytical Approach: The Integrated Analytical Framework (IAF)

As I have stated earlier, my objective for carrying out the literature review was to create a framework that I could use to evaluate the characteristics of the feedback data that I gathered during this research. I call this framework the Integrated Analytical Framework (IAF).

My approach to synthesising the literature into a framework involved a series of stages, which mirrored, to some extent, the thematic analysis steps outlined by Braun & Clarke (2006). Once the relevant literature was identified through the search process, I read each abstract to check for relevance, and then retained those that fitted the scope of the project. I then read each of these articles, in effect undertaking a process of familiarisation with the literature, and identified the main outcomes. These outcomes were noted, which was a form of initial encoding (e.g. ‘Detail’, ‘Quantity’, ‘Communication Mode’, etc.). Part way through the literature review process I was able to cluster some codes into nascent themes, which I called ‘Factors’, based on common elements. For example I decided that the ‘Detail’ and ‘Quantity’ codes could fit into an overarching grouping and become ‘Subcomponents’ of the ‘Content’ factor. Through this repeated process, which was akin to a theme generation process, I was able to define my initial framework.

This process then allowed me to look closely at the Subcomponents and therefore to construct a detailed view of each Factor. I considered the literature outcomes for each factor

subcomponent and constructed a bipolar criterion that could encapsulate the literature findings. I called these criteria 'Qualities' as they captured the fine level detail of each factor. For example, the literature dealing with the 'Detail' outcomes could be structured around whether the feedback content was 'General' or 'Specific' in nature. These fine level 'Quality' descriptors allowed me to gain a holistic picture of the feedback literature (e.g. how many factors may be considered to describe the feedback phenomenon), as well as to allow me to gain a picture of what the literature conveyed about feedback effectiveness (e.g. the relative weight of literature suggesting that 'Specific' versus 'General' feedback was most effective).

As I moved through the literature review process I was able to re-evaluate each factor or subcomponent as necessary, until I reached saturation and I was satisfied that no new factors or subcomponents of the framework could be generated (Table 2.4).

Table 2.4: IAF Factor and Subcomponent Structure

	Factor	**Sub-component**	**Qualities**		
Core	1 Language Use	a) Clarity of Communication b) Discourse Characteristics	Clarity Cohesion	←→ ←→	Vagueness Distancing
	2 Content	a) Detail b) Quantity	General Restricted	←→ ←→	Specific Elaborated
	3 Timing	a) Temporality b) Frequency	Immediate Often	←→ ←→	Delayed Infrequent
	4 Form	a) Mode b) Interaction Condition	Oral Collocated	←→ ←→	Written Remote
Socio-cultural	5 Source	a) Number b) Trust	Singular Respect	←→ ←→	Multiple Distrust
	6 Emotion	a) Evaluation	Positive	←→	Negative
	7 Recipient	a) Feedback Seeking b) Confidence c) Locus of Control	Avoidance Secure Self	←→ ← → ← →	Seeking Insecure Other
	8 Knowledge	a) Codification	Codified	← →	Tacit
	9 Feedback Giving	a) Explicitness	Explicit	←→	Implicit
	10 Pedagogy	a) Model of Learning	Transmission	←→	Construction

It emerged quite early in the review process that I could differentiate between some generic and some context-specific feedback features, and this led me to consider how I needed to encompass a broad range of conceptualisations of feedback within my analytical framework.

My consideration of findings from some previous meta-reviews (Hattie & Timperley, 2007; Kluger & DeNisi, 1996) enabled me to identify the final organising principle of the IAF, and this was the separation of 'core' and 'socio-cultural' factors. The inclusion of core and socio-cultural factors represented the final organising principle of the IAF. For me, core factors are

those that are to an extent context-free or standalone (Language Use, Content, Timing, and Form), whilst socio-cultural factors are necessarily context-specific or tied to particular situation (e.g. features relating to a recipient or the feedback giver).

I was anxious that a framework that attends primarily to general, or standalone, feedback features could implicate a behaviourist paradigm, a limitation also noted by Sadler (1989). I was concerned that a focus on decontextualized elements (such as information quality or timing) might reduce feedback study to being a simple phenomenon of stimulus and response. As a result, there might be an implication that the manipulation of these individual factors leads to learning changes, without adequate recognition that any influences might be mediated by important contextual or agent (e.g. recipient or interpersonal) factors. It is possible that focusing on feedback characteristics such as content and timing foregrounds the notion that feedback simply performs a transmission function in the learning process, and according to Nicol & Macfarlane-Dick (2006) this is a problematic assumption that underestimates the learning potential of feedback. This issue is discussed later, but it also hints at an important problem expressed by some commentators, who argue that feedback studies in the past have often been based on a weak conceptualisation of feedback that lacks a strong link with theory (Boud & Molloy, 2013a; Sadler, 1989).

It appears that feedback and behaviourist learning theory have a longstanding shared heritage which may be traced through two complementary historical developments. Recio Saucedo et al. (2013) highlight the way that the feedback concept shifted from its original roots in the physical sciences to the social scientific domain. For example, Recio Saucedo et al. (2013) point to the strong links between studies of feedback and cybernetics, citing as evidence the etymological roots of cybernetics in the Greek word *kubernētēs [κυβερνήτης]* (meaning steersman) to explain the tendency for such studies to over-focus on the regulatory feature of feedback. On the other hand, some commentators suggest that the prevalence of behaviourism in the development of educational psychology and testing in the early 20th Century has also had a profound influence on the notion of feedback. Kluger & DeNisi (1996) argue that the influence of Thorndike's law of effect (Thorndike, 1927), and its subsequent impact on educational psychology in the US, had a major impact on the developing science of educational assessment, and specifically on the development of the influential US Educational Testing Service (Bennett, 2007). This interpretation is supported by a review of feedback literature which outlines how the psychometric testing tradition has

been a major influence on subsequent models of feedback practice in education (Mory, 2004), and that this stretches beyond the US educational testing context (Butler Shay, 2004).

I argue that a weak conceptualisation of feedback has at least two consequences for a study of feedback. Firstly, it leads to a reliance on behaviourist theories that make it difficult to explain how feedback influences learning within complex environments. Secondly, a weak conceptualisation of feedback can undermine any ambitions to develop effective learning feedback advice and practice. As Boud & Molloy (2013b) note, 'we [currently] do not have a sufficiently secure idea of what feedback is for us to consistently use it effectively' (Boud & Molloy, 2013b, p. 1).

As a result of these concerns, I wanted my framework to incorporate both general or standalone features (core factors) alongside other context specific factors. The inclusion of these additional factors allows me to take into consideration features of feedback practice that pertain to the individuals and the contexts in which feedback is employed. To do this I included feedback features that were identified in a series of studies that foreground the situated nature of feedback activity (Brown & Glover, 2006; Hyland, 2001; May, 2013; Paul, Gilbert, & Remedio, 2013; Whitelock, Watt, Raw, & Moreale, 2003). These features include the nature of the knowledge being discussed, the characteristics of the participants, and the features of the institution within which feedback occurs. Drawing on the terminology of Paul et al. (2013), these are considered to be 'socio-cultural' factors. This is because they relate specifically to the local context of feedback discourse and highlight the importance of taking into account immediate contextual features when considering how feedback works.

In summary, I have outlined the background to the development of my IAF, which I developed with the purpose of evaluating the characteristics of the feedback data that I gathered during this research. In the next section I discuss each of the factors and subcomponents of the IAF in detail, before summarising how these relate in general to the study of feedback effectiveness.

2.4 Review Outcomes

2.4.1 Language Use

Two subcomponents were identified which related to the language use factor of the IAF. Research relating to the first subcomponent (clarity of communication) included research which explored variances in the implicit/vague or explicit/clear qualities of language use. Research relating to the second subcomponent (discourse characteristics) explored how language can be used to create either cohesion or fragmentation/distancing.

2.4.1.1 Clarity of Communication

Work-based feedback literature suggests that explicit and clear articulation is crucial for feedback message effectiveness. Johnson & Black (2012) found examiners prefer clear and explicit feedback messages from more senior examiners. It is argued that clear articulation is key to feedback effectiveness because message clarity enables legitimate interpretations to be constructed by the receivers of feedback (Vonderwell, 2003) and this minimises ambiguity on the part of the recipient (Johlke & Duhan, 2000). King (2016) reports that vague feedback attenuates the message force for the recipient. This is reinforced by a meta-review of literature carried out by Hutchins et al. (2013) and, interestingly, it has been observed that 'blunt' feedback is useful for dealing with underperformance in high stakes (i.e. clinical) training contexts (Bearman et al., 2013, p. 538).

A number of studies consider the nature of explicitness and clarity, centring on 'rationale-explanation' and 'being specific'. According to Nestel et al. (2013), discussing the clinical training context, effective feedback involves giving rationales in comments to novice clinicians in order to justify ratings given, to affirm excellent performance, to identify inadequate practice and to prompt learner reflection. This mirrors the findings of Ahmed et al. (2012) who argue that clinician feedback should highlight specific areas needing improvement. For Mengis & Eppler (2008), involving rationales in feedback allows feedback givers to engage in advocacy by providing data and explaining reasoning. Explanation of reasoning can also make learning expectations explicit through exemplar and model provision (Dennen, Aubteen Darabi, & Smith, 2007).

Specificity is also a recognised aspect of clear articulation. According to a work-based study by Bosley & Young (2006), specificity involves giving direct answers, explaining and elaborating. These actions are aspects of the cognitive exchange that allow participants to

engage in targeted critical analysis of performance with one another (Bosley & Young, 2006, p. 363). A number of commentators focus on specificity in terms of particular word use. Coninx, Kreijns, & Jochems (2013) found that the use of pre-specified, standardised terms in the feedback messages given to trainee teachers reduced the cognitive load required to simultaneously interpret the message and to complete the task at hand. Similarly, looking at language use in business contexts, Mengis & Eppler (2008) identify effective feedback as involving specific language that avoids euphemisms and focuses on the issues that matter the most.

This point mirrors those made by Huffaker (2010), Asmuß (2008), and Hyatt (2005), who note that remote feedback messages which use powerful language (e.g. imperative use to signal authority), and which avoid powerless cues (e.g. indirectness, hedging and passives that elide agency), are most persuasive and influential. Hedging has been defined as 'words whose job is to make things more or less fuzzy' (G. Lakoff, 1973, p. 175). At the same time, some research findings suggest that overly specific language can also undermine the potential of feedback to encourage learning. Research with novice clinicians and trainee teachers has found that the use of questions in feedback can prompt critical reflection (Hudson, 2014) and that the use of non-direct questions in particular can engage learners in their own learning through encouraging attentive listening (Ahmed et al., 2012).

2.4.1.2 Discourse Characteristics

Although it has been noted that studies of organisational and professional learning 'lack extensive descriptive accounts on the micro processes of conversations' (Mengis & Eppler, 2008, p.1293), there is a growing literature that looks closely at communication within workplace contexts. Some studies suggest that a defining feature of the feedback genre is that it involves managing bad news (Yelland, 2011; King et al. 2008). Looking at management discourse literature, Sussman & Sproull (1999) identify four common strategies that are used for delivering bad news: 'straight talk'; using 'positive politeness' (emphasising empathy and common ground features between participants or 'sugar coating' the message); 'negative politeness' (creating a respectful distance between participants); and 'omitting' bad news. Whichever way this bad news is (or is not) communicated, it appears that the priority for the feedback giver is to maintain coherence within the message, to avoid fragmentation of common understanding and to maintain an on-going feedback discourse to ensure improvement.

Coherence is a macro-level feature that enables meaning to be constructed (Herring, 1999), and according to Örnberg Berglund (2009) it is subjective and is something that is perceived by participants. For Dickey, Wasko, Chudoba, & Bennett Thatcher (2006) coherence is the relation of textual units to other textual units. Coherence building can be supported through the use of cohesive devices (Halliday & Hasan, 1976) and ensures that participants establish common ground with each other and share common expectations. Common ground is a concept borrowed from Clark (1996), and according to Hancock & Dunham (2001) it involves information that emerges in conversation and implicates knowledge that participants believe they hold in common with each other. This is an area that is discussed in a later section (Factor 10: *Pedagogy*).

Cohesive devices include reference use (e.g. pronoun use or relation of a text to other shared non-literary events) (Geertz, 1983); substitution (e.g. word replacement; ellipsis; conjunction through linking words); lexical cohesion (e.g. repetition); and, routinisation (e.g. the use of stock phrases and idioms that live in the interaction) (Garrod & Pickering, 2004). Severinson Eklundh (2010) adds quoting to the list of cohesive devices used in remote feedback messages. Quoting controls the discourse by limiting other potential directions for discourse to travel, but also engenders a pragmatic conflict. Reading the additional quoted content leads to supplementary processing but thereby includes further contextual data, suggesting that the feedback giver must strike a balance when deciding whether to include quotes in the text. Tuck (2012) observes that assessment feedback that employs informal markers can encourage reciprocity and dialogue. Examples include the use of repeated exclamation marks, underlining, and italics as signals for emphasis.

Literature also suggests that coherence can be supported through message structuring, e.g. through the maintenance of adjacency pair structures (Schegloff, 1992). Adjacency pairs are two-part exchanges in which the second utterance is functionally dependent on the first, and can be witnessed in conventional greetings, invitations and requests. Although these patterns are usually found in spoken discourse, the example of an adjacency pair structure shown below demonstrates how they are also evident in remote email discourse. In this extract from an examiner and Team Leader [TL] email interaction, taken as an illustration from my

research data[5], the TL mirrors the examiner's greeting and then thanks the examiner for the information that they have submitted to the TL.

From: Examiner: 08/07/2014 15:09:57

Hello *TL name*,

I've marked another batch - would you like to check before submission?

Examiner name

> ---
> ***>From: TL: 09/07/2014 08:49***
>
> Hi *Examiner name*
>
> Thank you for advising me that there is another batch to be reviewed.
>
> The feedback is as follows:
> …
> ---
> [E3 9.7 0849]

The example above also demonstrates the relation of dependence between the two parts of an adjacency pair, where the first part leads to expectation in a second part and establishes conditional relevance across the discourse structure (Severinson Eklundh, 2010). For Herring (1999), turn adjacency gives communication context and allows referential links to be used, capitalising further on the effectiveness of cohesive devices, mentioned above. According to Lapadat (2007), discursive coherence includes the use of reference devices to support backward reference; endogenous reference; and forward reference. This reinforces the point that feedback is a sequential rather than isolated process (Archer, 2010).

Literature suggests that the nature of feedback implicates at least two problems, with consequences for those who are responsible for giving feedback. The first problem relates to the way that the inclusion of negative information in feedback communication influences feedback message construction, including both the words used and the linguistic structures employed. Ackerman & Gross (2010) argue that the inclusion of negative information in feedback can increase the transactional distance between the feedback giver and the recipient. Others suggest that negative information can also contribute to a deteriorating situation

[5] This and other examiner feedback extracts in this chapter are taken from my research data.

(Yelland, 2011), as well as creating potential problems for on-going relationship maintenance (Kerssen-Griep & Witt, 2015). As a result, Chur-Hansen & Mclean (2006) note that providing negative feedback is a demanding skill, with a tutor needing to consider interpersonal issues when drafting feedback messages. These issues have led Geddes & Linnehan (1996) to observe that feedback that contains negative information is more complex than feedback communication that contains positive information. This complexity relates to the indirect and ambiguous messaging that more commonly correlates with negative feedback; this can adversely influence recipients' interpretations and ability to act on the feedback.

At this juncture, attention should be drawn to the work of theorists who suggest that the presence of negative information in social interaction necessarily embroils issues of face management and politeness (e.g. Goffman, 1967). The concepts of face and politeness are useful as they help to describe a tension that appears in the feedback literature around how to manage social cohesion and professional distance. For Morand (2000), criticising, disagreeing and interrupting represent potential face threat, and politeness can minimise this threat. Politeness may be of a positive or a negative variety, and each has different characteristics. Morand (2000) suggests that the use of negative politeness tactics reinforces distance in remote interactions and acts as a form of 'social brake' (Culpeper, 2001, p. 244). These negative tactics involve the judicious use of words to construct messages that include apology, verbal hedging, past tense use, honorific terms use (e.g. formal title), formal word use, and making the feedback more impersonal through avoiding the pronoun forms 'I' and 'you'.

On the other hand, positive politeness tactics reinforce commonality and act as a kind of social accelerator (Brown & Levinson, 1987, p. 103). These positive tactics include comments on admiration and the use of 'in group' speech forms (e.g. the use of ellipsis and the inclusive pronoun form 'we'). Morand (2000) found that positive politeness is used more by subordinate participants 'when performing a conversational act that may infringe or otherwise threaten a superior's face' (Morand, 2000, p. 244).

Another problem for feedback construction relates to the way that hierarchic professional roles can interact with feedback-giving. Focusing on clinician training feedback, Henderson, Ferguson-Smith, & Johnson (2005) note that asymmetrically structured learning tends towards unidirectional feedback-giving. For Archer (2010) these environments tend to focus

attention on the teacher, rather than learner, and pose a challenge to feedback givers about how to take into account the psychosocial needs of the recipient whilst also maintaining professional standards.

The problems outlined above have a number of consequences on the way that feedback communication is constructed. In line with the observations of Geddes & Linnehan (1996), Sussman & Sproull (1999) draw attention to the finding in management research of a tendency for information givers to distort feedback messages, particularly when giving bad news. They argue that this is due to the psychological anxiety related to the anticipated reaction of the message recipient. Moreover, this phenomenon can be related to the widely reported *MUM effect* which describes how, 'in general, individuals display greater reluctance to share bad news as compared to good news' (Dibble & Levine, 2010, p. 1).

Feedback construction can therefore be seen as involving the manipulation of communication to influence on-going interaction management. Yelland (2011) argues that feedback construction can essentially be seen as involving a set of moves around managing negative information. For example, the positioning of positive information prior to negative information is a structure that is commonly found within feedback messages (e.g. Chur-Hansen & McLean, 2006), and Yelland (2011) argues that this relates to the way that it helps to reinforce esteem and reduces anxiety. This form of structure**,** with positive information being foregrounded, might cue that there is an underpinning trust to the message and that any following negative comments are for the benefit of the receiver and not hostile in intent. Trees, Kerssen-Griep, & Hess (2009) and Yelland (2011) suggest that this also increases the likelihood that the recipient will consider the feedback to be useful and take the information into consideration. This argument is also reinforced by literature from some workplace feedback studies. Sias, Gallagher, Kopaneva, & Pedersen (2012) found that feedback givers can create closeness through emphasising personal over professional issues where there is a fear that distance is growing. Sias et al. (2012) also note that closeness can be managed through circumspection (e.g. avoiding uncomfortable/negative information); and deception/distortion (e.g. withholding information). They argue that this sort of approach can lead to the feedback recipient being treated as a 'whole person' rather than a 'work-role occupant'.

2.4.2 Content

Two subcomponents were identified which related to the content factor of the IAF. Research relating to the first subcomponent (Detail) included literature that explored variances in the level of the generality or the specificity of the feedback message. Research relating to the second subcomponent (Quantity) considered whether restricted or whether elaborated messages influenced effectiveness.

2.4.2.1 Detail

Feedback messages can involve information at different levels of detail and it seems that effective feedback provision involves decisions being made about the appropriate level of detail to be included in a message. Studies have suggested that goal specific feedback is most effective (Coninx et al., 2013; Thurlings et al., 2012). This focus on goals implies that the participants share a common understanding of the goal of the task, and that feedback coheres with the personal ambitions of the recipient (Archer, 2010).

There is evidence to suggest that feedback to examiners is most effective where it is individual to the recipient rather than being generic (Johnson & Black, 2012a). This specificity can also be useful if it includes linkages between and across the feedback recipient's performances (Chetwynd & Dobbyn, 2011). This individual focus is potentially effective because the feedback comments relate specifically to an examiner's performance, rather than restating the terminology of the mark scheme in an unmediated form. This point also relates to the work of Knoch, Read, & von Randow (2007), who note that the use of generic feedback comments relies on examiners decoding them to relate them to their own practice. Knoch, Read, & von Randow (2007) argue that this is less effective as it relies on examiners being able to notice and remediate their own biases, and that this task is more effectively attained through the involvement of external input.

Others note that detailed feedback, conveying information that is specifically targeted to the task, is most effective for learners in professional contexts such as clinical and teaching settings (Ahmed et al., 2012; Archer, 2010; Barton & Wolery, 2007). These details include consistent, transparent model answers or exemplars, or annotated comments on models (Li & De Luca, 2014). In professional examination contexts, the articulation of marking rationales to examiners in feedback was found to be effective (Greatorex & Bell, 2008). This point is reinforced by a number of studies that look at ineffective feedback. In many of these contexts ineffectiveness is connected with the inclusion of 'generally positive' comments, lack of

exemplars (Taylor, 2011), and the omission of an explanation around an erroneous performance (Li & De Luca, 2014; Segoe, 2013).

2.4.2.2 Quantity

The second characteristic of content that relates to feedback effectiveness is quantity. The literature around this aspect appears to involve a number of tensions around time and around amount/relevance.

Reporting on feedback-giving in health and social care contexts, Dearnley, Taylor, Laxton, Rinomhota, & Nkosana-Nyawata (2013) highlight that one of the main challenges for professional feedback giving centres on time pressures. Giving feedback is time consuming and this adds to feelings of frustration for feedback givers (Harms & Roebuck, 2010). Time pressures are also reported to be responsible for undermining ambitions of feedback givers who seek to offer elaborated, extensive feedback (Chetwynd & Dobbyn, 2011). This problem leads Whitelock, Gilbert, & Gale (2011) to pose the question about how it might be possible to improve feedback-giving without increasing workload. A potential solution they suggest is for the development of audio feedback systems, although this can also have implications for the character of the communication (Whittaker, 2003). This issue is discussed later (Factor 4: *Form*).

Issues around feedback amount and relevance are a second area of challenge. There are mixed messages in the literature about how feedback quantity influences feedback effectiveness. Dennen et al. (2007) report that students in remote settings perceive timely feedback to be more important than feedback extent. They also suggest that expert over-communication can lead to reduced learner interaction. This finding is supported by Ackerman & Gross (2010), who report that applied learners in marketing training prefer not to receive extensive feedback comments because this leads to a perception that the tutor has a negative impression of them as a learner.

These findings contrast with other study findings which suggest that learner understanding in some applied contexts, i.e. engineering, is better when feedback includes explanatory comments and reasoning (Taylor, 2011). Similarly, elaborated feedback, which includes verification and guidance has been found to be effective for scientific knowledge learning in remote contexts (Smits, Boon, Sluijsmans, & van Gog, 2008), and for improving task comprehension in remote language learning contexts (Murphy, 2010).

2.4.3 Timing

Two subcomponents were identified which related to the timing factor of the IAF. Research relating to the first subcomponent (Temporality) included literature that explored variances in the immediacy or delay of feedback messages. Research relating to the second subcomponent (Frequency) explored elements relating to how frequent or infrequent communication can influence feedback effectiveness.

2.4.3.1 Temporality

Literature suggests that communication temporality (i.e. immediacy and delay) factors can be manipulated to influence the effectiveness of feedback. Learners commonly report that they perceive timely and on-demand feedback to be beneficial for learning (Boling, Hough, Krinsky, Saleem, & Stevens, 2012; Dennen et al., 2007; Owens, Hardcastle, & Richardson, 2009)

A number of studies define effective feedback as having immediacy (Johnson & Black, 2012b; Chur-Hansen & Mclean, 2006; Ahmed et al., 2012; Burke et al., 2009; Hatzipanagos & Warburton, 2009; Mathieson, 2012; Li & De Luca, 2014). This impact has been reported across a variety of learning contexts, including clinical training environments (Archer, 2010), pre-service teaching environments (Barton & Wolery, 2007), and in a meta-review of computer-based feedback studies (van der Kleij et al., 2015).

The surveyed literature contains a number of reasons why immediate feedback might be considered particularly effective for learning. For Panahi, Birjandi, & Azabdaftari (2013), this effectiveness relates to the way that feedback can be incorporated within on-going learning processes. Münzer & Holmer (2009) link this to theory and highlight how media synchronicity theory (similar to Clark & Brennan, 1991) incorporates feedback (i.e. response) immediacy among its five media dimensions.

A proposed consequence of including immediate feedback is to lower the mental effort required for discourse participants whilst they try to maintain cohesion of meaning throughout a discussion. In other words, immediacy of feedback response is important for linguistic coherence and facilitating the acquisition of mutual understanding. Similarly, teachers in a study using delayed or asynchronous communication reported perceptions of disjoint and a need for immediate communication to reduce frustration levels (Winter & McGhie-Richmond, 2005).

Immediate feedback also appears to have a social dimension. Rovai (2000) suggests that immediate feedback reduces isolation, increases social presence and increases learner motivation. In the same way, Stein, Wanstreet, Slagle, Trinko, & Lutz (2013) found that immediate feedback can increase cognitive presence which aids meaning making through increasing the exploration of ideas and integration of information.

On the other hand, delayed feedback is considered effective for supporting different learning functions than immediate feedback (van der Kleij et al., 2015). Archer (2010) notes that feedback that is given with a delay between performance and feedback is better for supporting the transfer of knowledge, whilst Barton & Wolery (2007) argue that delayed feedback is effective in reinforcing already learned skills.

2.4.3.2 Frequency

Frequency is the second element of quantity that can be manipulated to influence the effectiveness of feedback. Two studies suggest that professionals perceive regular and on-going feedback to be most effective for improving their performances (Chur-Hansen & Mclean, 2006; Johnson & Black, 2012b). A potential explanation for this might be that communication frequency leads to cognitive salience, defined as how readily or frequently something comes to mind (Wilson & Williams, 1977, p.985).

Despite these perceptions, some literature findings question the relative effectiveness of high frequency feedback. Lurie & Swaminathan (2009) found that the frequency of feedback given in a business simulation influenced the degree to which recipients processed and acted on that feedback. They found that high frequency feedback led participants to lose sight of the bigger picture, reduced their ability to make comparisons of individual cases, and led to the participants paying most attention to the most recent information that they received. This led to the conclusion that high frequency feedback can diminish recipients' decision-making performances. This point is also mirrored in the findings of Morgan & Symon (2002), who note that information overload leads to ignored messages in remote workplace communication interactions. This message is reinforced by Cook et al. (1999), who observe that frequent feedback supports task completion but not long term learning in professional contexts.

Finally, there is evidence in the literature that feedback frequency impacts effectiveness because it encourages relationship building. In workplace studies, Gregory & Levy (2012) report that the frequency of feedback interactions that include positive information leads to

relationship building. Wilson & Williams (1977) take this further by noting that the combination of frequent communication involving the transfer of deep information (i.e. involving personal and personally significant information) is most effective. They argue that this combination of information and frequency allows participants to envision each other's perspective and 'develop mental images of other's work spaces, workloads and work habits' (Wilson & Williams, 1977, p. 985). This mutual identification allows the creation of common ground, reducing potential uncertainty with regards to the attributes of remote participants, and engendering positive attributes to others where data is lacking. It is also suggested that frequent communication from management to subordinates can increase subordinates' professional attachment and belonging as they are party to important information about work (Morgan & Symon, 2002), and this increases the likelihood that they will be influenced by feedback messages (Veloski, Boex, Grasberger, Evans, & Wolfson, 2006).

2.4.4 Form

Two subcomponents that related to the form factor of the IAF were identified. Research relating to whether messages were communicated in oral or written forms was considered in the first subcomponent (Mode). Research relating to the second subcomponent (Interaction Condition) looked at whether collocated (e.g. face-to-face) or remote communication influenced feedback effectiveness.

2.4.4.1 Mode

There is evidence in the literature that feedback givers are aware of the relative affordances of different communication modes and this can inform their choices about which mode to communicate through. For example Golden, Beauclair, & Sussman (1992) report that remote learners' choice of email communication related to their perceptions of the mode's qualities. In a workplace study, Morgan & Symon (2002) found that the type of message being communicated influenced the choice of communication mode, with the preference being for the face-to-face mode for communication about personal issues.

A number of studies suggest that modes that allow multiple streams of information are favoured over single stream channels (e.g. Johnson & Black, 2012b; Mathisen, 2012). These studies are supported by media synchronicity theory which suggests that s*ymbol variety* (the ability to provide information in various formats such as verbal or non-verbal forms) helps to underpin effective communication (Münzer & Holmer, 2009).

Video communication commonly employs dual channel functionality (i.e. visual and audio channels). Bietz (2008) observes that audio-visual communication carries more information than text, by as much as four times according to Stannard (2008). Moreover, audio-visual communication has been found to enhance participants' development of a shared view through the combination of comments and shared views of objects (Stannard, 2008).

Where single streams are discussed, studies suggest that audio, video and text modes can support different functions. In general, audio feedback is perceived as more detailed and easier to understand than text-based feedback. In a nurse training context, Gould & Day (2013) found that audio feedback allowed academic language to be broken down through two-way discussion to enable clearer interpretation. Similarly, some remote learners perceive audio feedback to carry more nuance than written feedback (Mathieson, 2012). These points allude to two affordances of audio communication. Firstly, audio feedback takes less time to produce and includes more words than written feedback (King, McGugan, & Bunyan, 2008). Secondly, audio communication often accommodates two-way verbal interaction, allowing clarification and aiding message reception (Blair & McGinty, 2013).

On the other hand, audio communication has some observed drawbacks, which might be overcome by written communication. King, McGugan, & Bunyan (2008) observe that giving negative feedback verbally can be uncomfortable. This is supported by Derks & Bakker (2010), who report that bad news is more directly communicated by email. Written communication also provides a record of the interaction that can allow reprocessability (Dennis, Fuller, & Valacich, 2008), and encourage learner reflection through (a) examination and subsequent processing (Kim, Park, Yoon, & Jo, 2016; Mancilla, Polat, & Akcay, 2017; Münzer & Holmer, 2009), and (b) annotation use (Bridge & Appleyard, 2005). In addition, written text can be delivered without breaking the professional flow of an activity, encouraging later reflection (Barton & Wolery, 2007). Finally, written feedback via email can contain an implicit invitation for learners to engage in discussion (Barton & Wolery, 2007; Hatzipanagos & Warburton, 2009).

2.4.4.2 Interaction Condition

Despite claims that miscommunication is not a characteristic of technology but a characteristic of participants' gaps in shared understanding (Dickey et al., 2006), there is literature to suggest that communication mode does influence communication effectiveness. Collocated communication, where all parties are physically present, shares some of the

characteristics of dual channel communication and is considered to be effective for particular communication functions. Many of these reflections are made in comparison with computer-mediated communication (CMC), which is generally remote and largely text-based (e.g. email).

Literature suggests that collocated communication appears to have a positive impact on learning-related tasks. According to Hebert & Vorauer (2003), collocated communication leads to more accurate feedback being given on skills compared to CMC. Ellis, Goodyear, Prosser, & O'Hara (2006) report that students perceive it to be easier to reflect on their opinions and analyse their experiences in a collocated communication mode. Similarly, collocated communication is reported to be better for intellective tasks (i.e. problems where there is a demonstrable correct answer), or for complex learning tasks that require negotiation and convergent thinking (reaching consensus) (DeSanctis & Monge, 1998; Martins, Gilson, & Maynard, 2004).

A number of issues might influence the ability of participants to establish consensus through remote interaction. Mengis & Eppler (2008) citing the work of Donnellon, Gray, & Bougon (1986) in business learning contexts, suggest that this might be because face-to-face interaction allows participants to employ linguistic indirection (e.g. lack of precision) to express ambiguities around issues. Collocated interaction also allows participants to demonstrate their continued attention (Örnberg Berglund, 2009), affords immediate response which can minimise interpretive ambiguity (Herring, 1999), and can incorporate paralinguistic cues (e.g. gestures) which can reduce the chances of miscommunication and the need for participants to make assumptions about the tone of the communication (Kruger, Epley, Parker, & Ng, 2005; Winter & McGhie-Richmond, 2005). Johri (2012) highlights that judicious language choices are more difficult to make through CMC due to the restricted information that is available to participants about each other. Inaccurate impressions can reduce the ability of participants to predict and explain co-workers' actions. A consequence of this is that remote work-related feedback interaction tends to be more abstract and lacks specificity compared with collocated feedback activity (Liberman, Sagristano, & Trope, 2002, cited in Whitford & Moss, 2009).

On the other hand CMC is considered to be better than collocated communication for some tasks. Since CMC tends to contain fewer social cues, feedback that is given via this mode tends to be more negative (involving 'straight talk'), but also to be more honest (Sussman &

Sproull, 1999). It is argued that this phenomenon is explained by the observation that deliverers of negative feedback via voicemail are particularly uncomfortable with the interaction (compared with the message receivers, see Watts, 2007); that there is generally less concern with social aspects of communications in CMC (Hebert & Vorauer, 2003); and that the CMC discourse forum can be private and avoid concerns about the public loss of face (Santamaría-García, 2017). Finally, Herring (1999) argues that it is overly simplistic to suggest that CMC lacks coherence. It is recognised that 'back channelling' is a strategy that is used by participants in spoken discourse to display their continued attention and understanding (Lambertz, 2011). According to Fowler & Fowler (1974), back channelling might be a sound or gesture made to give continuity to a conversation by a person who is listening to another. This phenomenon is demonstrated in the extract below where the examiner and TL use the adverb 'yeah' in the midst of each other's utterance to display their engaged listenership (lines 011 and 015).

```
008   TL    Yah, I'm afraid I've got to send them back so
009         you'll have to do another set I'm afraid? And
010         then submit those before we can [be]
011   Ex                                    [Yeah]
012   TL    up and running. Now is there anything in
013         particular, you, you said something
014   Ex    I download some and I've marked some [I'll]
015   TL                                         [Yeah]
016   Ex    send you those when I've finished them

[E1 619.1200]
```

Herring (1999) goes on to argue that CMC can actually be 'advantageously incoherent' (Herring, 1999, p. 1), since back channelling can also exist in asynchronous email (e.g. through the inclusion of messages with no significant new content which signal listenership), and through the use of quotes as a form of linking to connect messages to previous content. This additional information is considered to be helpful since it gives participants evidence that they share continued attention (Cornelius & Boos, 2003). The extract below shows how a TL interjects their discourse (in bold font) into the body of an examiner's original query, linking their ideas to those of the examiner.

From: Examiner: 18/06/2014 12:08:35

Hi *TL name*,

Thank you for the feedback, I have amended the 2 you sent back to me.

2 queries:

ID 649581302 - Q1 g ii - MS says 'it' should be assumed to mean cyclohexane. Do they still need to have written cyclohexane somewhere in their answer to get the mark? I accepted 'It burns more effectively'.

I cannot find the comment re: 'assumed to be cyclohexane' in the mark scheme - it may have appeared in the practice scripts by the sound of it, and was incorrect if it did..

ID 649661411 - Q2b - do they get the mark even though 'curly' arrow is almost straight?

I am afraid so - possibly a little generous.

I will look through the other 5 and send over.

Thanks.

TL name

Thanks,
Examiner name

[E3 18.6 1456]

2.4.5 Source

Two subcomponents were identified which related to the source factor of the IAF. Research relating to the number of feedback sources was explored in the first subcomponent. Research relating to the second subcomponent explored how elements of trust might influence feedback effectiveness.

2.4.5.1 Number

There are mixed messages in the research literature about the effects of feedback source issues on feedback effectiveness. Some studies suggest that learners perceive access to multiple feedback sources to be beneficial to their learning. Bearman et al. (2013) and Burke, Marks-Maran, Ooms, Webb, & Cooper (2009) highlight how multiple information sources

can aid learner self-recognition of underperformance in clinical and other applied learning contexts. Even in asymmetrical learning environments, there is research evidence that peer feedback is valued by learners. According to Segoe (2013), trainee teachers value mutual feedback from other learners and report that this increases levels of rapport and community.

In contrast, Bosley & Young (2006) report that single-stream (i.e. 1:1) dialogues between experts and learners are most effective for learning. This is because such discourse focuses mutual engagement around a process of targeted critical analysis and knowledge construction. This argument links with 'parallelism', one of the five dimensions of media synchronicity theory (Dennis & Valacich, 1999). Parallelism describes the number of simultaneous conversations within a communication stream, and theory suggests that a larger numbers of simultaneous conversations can increase incoherence (Münzer & Holmer, 2009).

2.4.5.2 Trust

There is evidence in the literature that issues of source trust and distrust also influence the effectiveness of feedback. In this area there appear to be three tensions. Firstly, trust is often considered to underpin participants' learning relationships. According to Gregory & Levy (2012), workplace feedback environments are considered to be strong and positive where, among other things, trust and empathy displays in feedback to work colleagues help to develop effective relationships between participants.

Gregory & Levy (2012) also note that the frequency of positive feedback interactions leads to relationship building in the workplace. Since the inclusion of positive information in feedback messages tends to coincide with the notion of face mitigation, it has been observed that recipients of positive feedback tend to perceive that the feedback giver is more credible (Trees et al., 2009).

It is also suggested that frequent communication from management to subordinates can increase subordinates' professional attachment and belonging as they are party to important information about work (Morgan & Symon, 2002), and increases the likelihood that they will be influenced by feedback messages (Veloski et al., 2006). Finally, where trust exists between communication participants there are reductions in perceived uncertainty in social exchange (Jarvenpaa & Leidner, 1998). This is reinforced in a meta-review of literature carried out by Hutchins et al. (2013).

The second issue that is identified in the literature is that communication influences trust, with the inference being that feedback giving can positively contribute to trust building. Morgan & Symon (2002) note that open workplace communication leads to trust, although they acknowledge that this is difficult to construct remotely. This point coheres with arguments drawing on media richness theory, which suggests that team trust (and performance) is improved when participants have access to more information about each other. Martins et al. (2004) observe that this often includes social and informal (i.e. non work-specific) information about other work colleagues.

Another challenge to trust building that is noted in the literature is the links between poor linguistic ability and unclear communication (Huffaker, 2010). This finding overlaps with Factor 1 of the IAF (*Language Use*) and points to the idea that poor communication skills reduce propinquity. Propinquity refers to the presence, nearness, or proximity of the feedback giver as perceived by a message recipient (Korzenny, 1978). Furthermore, this poor communication can undermine the recipient's perception of the feedback giver's credibility and reduce feedback influence. This point is picked up by Segoe (2013), who points out that ineffective feedback tends to be disregarded.

2.4.6 Emotion

Research relating to whether evaluative comments and the communication of positive or negative feedback information influenced effectiveness is explored in this factor. The research also considers the encouraging or demotivating elements of such feedback information.

2.4.6.1 Evaluation

Emotion is important to consider in relation to feedback effectiveness because it mediates learners' approaches to learning (Dowden, Pittaway, Yost, & McCarthy, 2013). Feedback has a well-known emotive impact on recipients and this can influence learning in general (Hattie & Timperley, 2007). This phenomenon is also recognised in applied and remote learning contexts (Furnborough & Truman, 2009; Gould & Day, 2013).

There is evidence that the inclusion of positive information in feedback has a beneficial impact on learners. For example, Mathieson (2012) and Stracke & Kumar (2010) report that effective remote feedback includes information about what a learner has done well and what they need to improve, and that expressive, praise-based feedback has the highest impact on

learners. It has been noted earlier that negative feedback might increase the transactional distance between feedback communicants (Ackerman & Gross, 2010). On the other hand, Thurlings et al. (2012) note that there appears not to be a definitive understanding in the research literature about the emotive influence of positive and negative feedback content on remote and/or professional learners.

According to a work-based review by Derks & Bakker (2010), it might be anticipated that positive feedback would have a beneficial impact on learning, as motivation is one of the two key purposes of feedback-giving (with the other being offering critical and objective evaluation on performance). Mathisen (2012) picks up on this theme, arguing that feedback is essentially motivational for learning and encourages integration of the learner into a professional community.

2.4.7 Recipient

Three subcomponents were identified which related to the recipient factor of the IAF. The first subcomponent (Feedback Seeking) considered research relating to whether recipients' feedback seeking or avoidance tendencies influenced feedback effectiveness. Research relating to the second subcomponent (Confidence) explored elements relating to how recipients' confidence levels can influence message interpretation. Research relating to the third subcomponent (Locus of Control) considered how recipients' perceptions of whether they are able to influence performance improvement could also influence feedback effectiveness.

2.4.7.1 Feedback Seeking

There is evidence in the literature that personal attributes influence the reception of feedback. For example, Ajjawi & Boud (2017) state that the effects of feedback 'are not realised unilaterally… [or] fully determined by the nature of inputs, but by what [the recipients] bring to them' (Ajjawi & Boud, 2017, p. 253). This point is reinforced by Fransen & Hoeven (2013), who report that negative feedback is better received when it is congruent with the 'regulatory fit' of the recipient.

Regulatory fit describes the extent to which the goal of an activity matches the expectations of the participant. This implies that feedback givers need to be aware of the expectations that learners bring to learning activities. Despite this, Gregory & Levy (2012) observe that workplace learners' orientation to feedback is little understood. Orientation is receptivity to

feedback and describes the likeliness of a learner to seek, value, and process feedback. Greatorex & Bell (2008) found some examiners were more apt to attend to feedback than others. This is also mirrored in the management literature where Linderbaum & Levy (2010) note that recipients differ in their responses to feedback.

Feedback seeking might be a key attribute for successful learners. For Linderbaum & Levy (2010) and Bosley & Young (2006), feedback seeking behaviour actively shapes the feedback environment by instigating learning discourse based on the learner's current level of understanding. Blair & McGinty (2013) suggest that this instigation attribute might relate to the personal characteristics of a particular learner. For example, Stracke & Kumar (2010) suggest that learners who have a high degree of self-regulation are likely to seek feedback. On the other hand, Whitaker, Dahling, & Levy (2007) argue that seeking might also be an environmental attribute, with learner perceptions of a positive supervisor/co-worker relationship increasing feedback seeking.

2.4.7.2 Confidence

There are mixed findings in the literature about how feedback interacts with personal attributes. Furnborough & Truman (2009) observe that learner confidence is a precursor to their attitude to feedback. It is reported that some learners prefer personalised, individual feedback. Reporting on learners in an asymmetrical academic learning context, Blair & McGinty (2013) suggest that having adequate time for discussion and interaction between tutors and learners helps to develop learners' self-confidence.

It is also possible that confidence interacts with feedback communication mode, with depersonalised feedback being most effective for some remote learners. For example, AbuSeileek & Qatawneh (2013) report that shy or anxious learners benefit from asynchronous, de-personalised interactions.

2.4.7.3 Locus of Control

A final attribute that appears to influence the degree to which feedback might be effective relates to the locus of learning control. Engagement with feedback might be greatest where the locus of learning control is internal to the learner (rather than being external to the learner). This is because the learner believes that they can influence their own learning through action. Ajjawi, Schofield, McAleer, & Walker (2013) observe that clinical learning is effective when learners actively engage with feedback. This is mirrored by Bietz (2008), who

notes that remote feedback is most effective when the recipient possesses a sense of self efficacy in the learning process.

One explanation for why control locus might be important is because the learners' own self-perception can influence the interpretations they construct around feedback messages. Again, in the clinical learning context, Bearman et al. (2013) argue that ownership of feedback by the learner is important, otherwise there can be a tendency for a learner to discount information that does not fit their own self-evaluation.

2.4.8 Knowledge

Research literature which related to the nature of the knowledge in feedback communication was explored in this factor. This included research relating to how the form of knowledge being fed back on (i.e. either codified or tacit) influenced issues of feedback effectiveness.

2.4.8.1 Codification

The nature of different types of professional knowledge is discussed in the literature. This is important to reflect on since these form an aspect of the context within which feedback practice exists. Billett (1993) outlines a model of workplace learning that comprises of three forms of knowledge: propositional or conceptual knowledge (knowing about/knowing that); procedural knowledge (knowing how); and dispositional knowledge (important values and dispositions). This model highlights that knowledge has context specific elements and that it can either be tacit or can more easily be codified. Moreover, it shares some overlaps with Bernstein's (2000) concept of vertical/hierarchic knowledge structures (which are well defined and can be articulated, e.g. maths, science), and horizontal knowledge structures (which are relatively fragmented, with weak integration of propositions, and contain highly tacit elements, e.g. social science).

The underpinning message of these models is that knowledge has context specific character (Gasson, 2005). The models can also be used to reinforce the idea that feedback has a role in helping participants to learn through making sense of such context-specific knowledge. According to Shalem & Slonimsky (2010), successful learning in horizontal knowledge structures is akin to the adoption of a gaze or position, and relies on experience and interaction with those possessing that gaze. This links with the idea of common knowledge, where participants in discourse invoke elements of their shared experience (e.g. subject

knowledge) as resources to ensure that a cumulative process of engagement is constructed (e.g. Mercer, 2008).

Formal structures of communication can also have an impact on the sense making processes of participants. Involving new participants in a professional genre allows insights into the ways of communicating, using language and thinking in a particular context (Gasson, 2005). In this way, practice and interaction is access to knowledge (Taylor, Dearnley, Laxton, Nkosana-Nyawata, & Rinomhota, 2012), with feedback correcting and guiding practice within the professional context (Barton & Wolery, 2007).

Literature suggests that effective feedback is able to contribute to professional learning via a number of mechanisms. Spafford et al. (2006) argue that feedback carries implicit tacit messages, such as through the amount of time devoted to dealing with specific content or the way that communication carries messages about the appropriate ways that participants should interact within a particular professional context. Feedback can also use explicit modelling (Schön, 1983) and verification (May, 2013) strategies to reinforce learning.

Feedback also has a codifying function. Gasson (2005) notes that business processes are largely tacit, but that collaboration between co-workers requires knowledge codification for exchange. Codification involves setting knowledge down in writing. This codification accords value to knowledge because, according to Stevenson (2001), affording value to unspecified knowledge is considered to be suspicious. This point reiterates that made by Polanyi (1969), who argues that codification is important because '…it allows us to talk about our knowledge, to inspect it, to assist in sharing it with others, to assist others in acquiring it, to engage in an appraisal of its coherence and utility…' (Polanyi, 1969, p. 648). For Edwards (2012), making knowledge visible is even more important where shared experience is limited, as this allows for the establishment of common ground for collaborative action.

The concept of expertise (which links with professional experience) also influences the ability to be able to give effective feedback. Polanyi (1969) argues that there is a split between 'what we know' and 'what we can say'. Experts are in a privileged position as they are able to relate particular pieces of information to the global objectives of a task (since it is these global elements which give particular pieces of information their meaning). In this way, 'knowing' can be said to be 'in practice' because it cannot be completely codified outside of a specific context (Stevenson, 2001).

2.4.9 Feedback Giving Process

Research literature which related to the feedback giving process was explored in this factor. This included research relating to how either explicit or implicit aspects of the feedback-giving process influence effectiveness.

2.4.9.1 Explicitness

Some of the reviewed literature unpacks the complexity of feedback. Ypsilandis (2002) outlines some of the decision-making that is inherent to feedback-giving which make it a complex activity. These decisions include considerations of quantity, length, content, frequency, style, and delivery method. King et al. (2008) highlight that feedback is generally a negative discourse, a finding that coheres with the results reported in Johnson (2013, 2016b), which needs to be skilfully managed at a social level to maintain recipient engagement (Mathieson, 2012).

Davis & Foster (2002) highlight that feedback-giving includes being able to look at feedback from the reader's point of view (one of the three elements of communicative competence, according to Canale & Swain, 1980), whilst Basturkmen et al. (2014) suggest that being able to express ideas with linguistic accuracy is a crucial skill for feedback givers. For Prins, Sluijsmans, & Kirschner (2006), a key skill for giving effective feedback is the ability to be able to translate information into language that is understood by the learner to improve their practice.

Consequently it can be claimed that relational expertise is involved in the learning process because it contributes to the quality of feedback-giving. According to Edwards (2012), relational expertise involves recognising what engrosses others and what knowledge matters. For Huffaker (2010), effective leaders in remote work reach out and engage in relationship development and management. This engagement leads to a social closeness, to the adoption of the leader's ideas and to a reduction in potential attrition (Owens et al., 2009). This links with the linguistic component of discourse mentioned earlier (Factor 1: *Language Use*).

Relational expertise can be inherent to feedback messages where a feedback giver is cognisant of the effect of a feedback message on the emotions of the recipient (Dowden et al., 2013). Moreover, the provision of individualised feedback, which Boling, Hough, Krinsky, Saleem, & Stevens (2012) claim can help to build a strong tutor-learner relationship, is also a component of relational expertise. It is possible that relational expertise is important for

learning feedback because positive relationships can lead to better alignment between participants' understandings. This accords with the work of Carpenter & Wisecarver (2004), who argue that building and maintaining relations is an element of interpersonal skill that allows people to achieve jobs together.

According to Cornelius & Boos (2003), participant familiarity 'fosters coherence, explicit references, and task orientation' (Cornelius & Boos, 2003, p. 170). Tseng & Kuo (2014) also report that the existence of closer attachments between teachers leads to a greater willingness for them to share resources. This could be important since these shared resources are elements of common ground that the discourse participants share. Finally, it has been observed that the use of humour can reinforce symmetrical and non-threatening social alignment so may be considered to be a component of relational expertise (Georgakopoulou, 2011).

There are a number of potential challenges to relationship building. It is possible that feedback givers' perceptions of others are inaccurate. In a review of work-based studies, Derks & Bakker (2010) report that it is common for those who instigate remote communication to use themselves as a reference point for their anticipation of how the message will be received by others. In remote workplace learning studies, it is argued that the physical distance between co-workers means that participants have less personal information about each other so that their remote impressions of each other lack complexity (Johri, 2012). Similarly, the development of a picture of other remote co-workers might be relatively slow because the development of shared understanding takes time to develop through a history of repeated actions (Dickey et al., 2006).

Some literature also highlights barriers to effective feedback-giving, including the pursuit of dual functions through feedback messages. Attempting to fulfil both 'work' and 'learning' functions can be a tension that is difficult to manage in same communication moves (Pryor & Woodward-Kron, 2014). It is also clear that feedback involves the management of social relations within the context of sharing negative content, which can also be problematic (Higgins, Hartley, & Skelton, 2001).

A number of studies suggest that characteristics relating to the feedback giver can undermine effective feedback-giving. These characteristics include egocentrism interfering with the ability to predict others' message reception (Kruger et al., 2005), and individually held attitudes and beliefs about how feedback influences learning (Hudson, 2014; May, 2013; Ros

i Solé & Truman, 2005; Tang & Harrison, 2011). Finally, there can be institutional pressures that act as barriers to effective feedback-giving. Chetwynd & Dobbyn (2011) observe that examiners spend most of their shared discussion time focusing on issues of awarding precise marks rather than propagating ideas for good feedback practice.

2.4.10 Pedagogy

Research literature which related to how the model of learning that underlies feedback (e.g. transmission or construction), and how it can influence feedback effectiveness, was considered in this factor.

2.4.10.1 Model of Learning

The reviewed literature supports the idea that discourse and interaction can encourage learning, with this discourse taking different forms, e.g. professional guidance (Hudson, 2014) and reflective discussion (Ellis et al., 2006). Commentators argue that learning discourse combines elements of skills content and elements of disciplinary norms and standards (Dannels, 2000), and specific types of judgements (Hyland, 2001) that highlight what it is 'to be' part of a professional community.

When considering why discourse and interaction encourage learning, literature suggests that communication helps to establish common ground and shapes and confirms mutual understanding (Cornelius & Boos, 2003; Kirschner & Lai, 2007). This is important because meaning construction relies on the development of common ground (or mutual knowledge) via discourse interaction (Beers, Boshuizen, Kirschner, & Gijselaers, 2007; DeSanctis & Monge, 1998). At the same time, literature suggests that common ground is developed through communication that brings together varying perspectives that lead to negotiated meanings (Espasa & Meneses, 2010). Considering this issue from a socio-cultural perspective, interaction expands learning in a joint activity as participants come to see a problem from another perspective (Edwards, 2012).

Interaction is also considered to influence professional learning as it can encourage identity formation (Raz & Fadlon, 2006). According to Ren, Kraut, & Kiesler (2007), group identity grows through shared communication, with shared task work and social interaction increasing bonding. An aspect of this bonding also involves the sharing of personal information through self-disclosure in interaction.

Much of the surveyed literature suggests that learning relies on the learner actively participating in meaning making (Furnborough & Truman, 2009). In this way learning is enabled through dialogue, with knowledge construction involving a number of elements such as sharing information, discovering inconsistencies, and agreeing on compromises (Paulus, 1999). Dialogue also allows learners to indicate areas of insecure understanding, which can open up opportunities for fruitful interactions that develop understandings (Li & De Luca, 2014).

Feedback can be looked at as contributing to a learning discourse, with skills or knowledge gap identification leading to learner action (Furnborough & Truman, 2009). In its simplest sense feedback is an indicator of what an expert looks for (Whitelock et al., 2003), and helps to bridge the gap between a learner's experience and making 'professional sense' of it (Ahmed et al., 2012, p.524). Through expressing the expectations of a professional community, feedback is a tool for induction into such a community (Basturkmen et al., 2014). In this way feedback synchronises the learner and the tutor, through articulating a degree of dissonance between the current and desired state of learner performance (Recio Saucedo et al., 2013; Adcroft, 2011). Literature also suggests that feedback enables self-regulation (Jordan, 2012; Espasa & Meneses, 2010) by triggering critical self-reflection (Hudson, 2014) and helping learners to realise why they are not meeting privileged criteria, thereby considering how they might develop more appropriate performances (Shalem & Slonimsky, 2010). In contrast, Harms & Roebuck (2010) argue that self-reflection should precede feedback, as this allows the learner to reflect on their thinking prior to confronting an expert opinion. Situating self-reflection prior to feedback can discourage any potential learner dependency that can undermine the development of learner agency.

Whilst the observations of Evans & Butler (1992) suggest that feedback can help to sustain a passive learner role in relation to the expert, mirroring conditions reported in some other asymmetrical feedback environments (e.g. Wang & Li, 2011), these findings contrast with the work of Smith (2003), which outlines how feedback can act as effective proximal guidance. Proximal guidance describes how feedback allows a learner access to, and an ability to construct, relevant conceptual knowledge. In this way, feedback interaction is a key to learning as it makes knowledge visible. Smith (2003) reports that most vocational learners prefer a proximal guidance learning approach as it encourages them to self-reflect; in this way feedback helps learners to become self-regulated learners (Stracke & Kumar, 2010).

The inclusion of modelling (i.e. the performance of a task with learner observation) is perceived by learners to be an effective strategy in the early stages of learning (Boling et al., 2012), and allows the context of a performance to be taken into consideration. Again, Johnson & Black (2012a) found that the most effective examiner feedback was considered to be that which they felt related to their own professional context.

2.5 Summary of the Effective Feedback Factors

Following on from the literature search, and the construction of the IAF to structure the outcomes of the search in a methodical way, I set out to evaluate the messages from the gathered literature regarding effective feedback.

As I outlined in Section 2.3, each factor that I identified, including any factor subcomponents, was organised as a bipolar continuum (labelled 'Qualities') to reflect the outcomes of the reviewed literature. For example, the literature outcomes relating to the 'Clarity of Communication' subcomponent of the 'Language Use' factor was considered in relation to the qualities of 'Clarity' or 'Vagueness'.

To evaluate the weight of evidence from my literature review I counted the number of articles that provided evidence in support of each 'Quality' in relation to feedback effectiveness (which I defined as being feedback that supported the alignment of perspectives in hierarchic or remote situations). I recognise that this is a crude measure of evidence quality but it gives an indication of where the majority of findings point in a particular direction.

Table 2.5 shows the IAF factors with a representation of the weight of literature reflecting each of the factor qualities (in terms of the number of supporting articles).

Table 2.5: IAF Factor Qualities and the Analysis of Effective Feedback

	Factor	Sub-component	Qualities	n[6]		n	
Core	1 Language Use	a) Clarity of Communication	**Clarity**	15	←→	2	Vagueness
		b) Discourse Characteristics	Cohesion	6	←→	15	**Distancing**
	2 Content	a) Detail	General	1	←→	12	**Specific**
		b) Quantity	Restricted	2	←→	4	Elaborated
	3 Timing	a) Temporality	**Immediate**	18	←→	3	Delayed
		b) Frequency	Often	6	←→	3	Infrequent
	4 Form	a) Mode	Oral	9	←→	9	Written
		b) Interaction Condition	**Collocated**	13	←→	4	Remote
Socio-cultural	5 Source	a) Number	Singular	3	←→	3	Multiple
		b) Trust	**Respect**	8	←→	0	Distrust
	6 Emotion	a) Evaluation	**Positive**	5	←→	2	Negative
	7 Recipient	a) Feedback Seeking	Avoidance	1	←→	4	**Seeking**
		b) Confidence	Secure	1	←→	1	Insecure
		c) Locus of Control	**Self**	3	←→	0	Other
	8 Knowledge	a) Codification	**Codified**	8	←→	2	Tacit
	9 Feedback Giving	a) Explicitness	Explicit	7	←→	16	**Implicit**
	10 Pedagogy	a) Model of Learning	Transmission	3	←→	21	**Construction**

This analysis showed that there is a clear difference in the weight of supporting evidence for the different feedback factor qualities. For this analysis I was also able to cluster the findings for some factors around common elements (e.g. literature around the *'Pedagogy'* and *'Recipient'* factors have findings that relate to the common concept of self-regulation). I present the following findings from the literature in order of magnitude.

The central, collated findings from the literature indicate that effective feedback involves:

- Discourse that seeks to support the construction of learning (*Pedagogy* factor), rather than merely transmitting important information. This message also shares links with literature that highlights how effective feedback has a participative dimension and an active recipient role (*Recipient* factor: Feedback Seeking; Locus of Control). These findings share a common socio-cultural perspective (Edwards, 2012) because they both recognise the important role of learner self-regulation in the learning process (Espasa & Meneses, 2010; Jordan, 2012; Stracke & Kumar, 2010);
- Giving feedback immediate to task completion (*Timing* factor: Temporality);
- Implicit elements that centre on relationship management – which also implicate participant familiarity levels (*Feedback Giving* factor), alongside explicit elements that seek to codify important information (*Knowledge* factor). These dual purposes represent a key component of feedback complexity;

[6] The number of articles that indicate that the quality contributes to effective feedback.

- Clear and specific language use (*Language Use* factor: Clarity of Communication; *Content* factor: Detail);
- Politeness and face management (*Language Use* factor: Distancing);
- Collocated interaction (*Mode* factor: Interaction Condition);
- The recipient having trust in the feedback giver (*Source* factor: Trust);
- Positive information feedback (*Emotion* factor).

The literature also suggested that there were mixed or inconclusive messages in relation to effectiveness and some feedback qualities. These included:

- The amount of feedback;
- The feedback mode (i.e. written or spoken communication);
- Feedback frequency;
- The number of feedback sources;
- Feedback recipient confidence levels.

Prior to discussing my study context I now use these findings to reflect on which factors are of specific interest for a study of remote or hierarchic working contexts.

Construction of learning through feedback and participant engagement: Remote or hierarchic working contexts can be challenging environments for crafting feedback that supports the construction of learning and engages active participation. This is because remote communication can include natural gaps in interaction that make it difficult to hold onto a train of thought. For example, in the case of email communication, gaps in interaction may coalesce around the affordance that allows people to send or open up a message when it's convenient for either of the participants to do so (Johnson, 2016a). This means that direct, real-time interaction can be a minimal element of some remote interaction (Herring, 1999). It is possible that in hierarchic discourse there can be a pressure for messages to be crafted that tend to either exclusively reflect the top-down views of the feedback-giver (Archer, 2010; Henderson, Ferguson-Smith, & Johnson, 2005). In addition, discourse may fail to reflect fully the views of a subordinate participant because of the fear of the potential personal costs of exposing weaknesses, and a fear of losing professional face (Goffman, 1967).

Immediate feedback: Remote or hierarchic working contexts can also be challenging environments for the delivery of immediate feedback. The gaps in remote feedback have been outlined above, and these gaps may be exacerbated by hierarchic working arrangements where the feedback giver has a number of people to communicate with in an intense period of

time. These commitments make it difficult to respond with feedback in real time or close to the moment of task completion.

Implicit elements of feedback giving: Hierarchic working relationships can present a challenge for feedback giving because the task of relationship management through communication may be considered to be less important than the explicit task of delivering important information. It is also possible that feedback givers may not recognise the learning needs of those they give feedback to because of a tendency to base information on their own expert perspective (Derks & Bakker, 2010; Kruger, Epley, Parker, & Ng, 2005).

Clear and specific language: Hierarchic communication may support the crafting of clear and specific feedback because the feedback giver has the authority to convey understandings that have been agreed at a senior examiner level. This singularity of message means that 'parallelism' is minimised and this can reduce potential incoherence (Münzer & Holmer, 2009). The remote nature of this communication also supports a degree of clarity because it tends to use a text-based format. Email has the affordance of allowing time to reflect on a message for clarity at the drafting stage. Email also acts as a record of interaction (Kim, Park, Yoon, & Jo, 2016; Mancilla, Polat, & Akcay, 2017). This record can aid interpretation through providing the reader with a resource that they can use to make anaphoric (backward) references and to check for sense (Herring, 1999; Lapadat, 2007; Severinson Eklundh, 2010). At the same time it needs to be acknowledged that the drafting of email is relatively time consuming (measured at a per-word rate compared with spoken communication) (King, McGugan, & Bunyan, 2008; Stannard, 2008), and lacks the paralinguistic elements of face-to-face communication (Münzer & Holmer, 2009). This issue may mean that messages may appear to be very literal (e.g. lacking some of the nuance of spoken discourse) (Mathieson, 2012), and reduce the likelihood that examiners will reply to feedback emails to ask for additional clarifications on all occasions. Where telephone communication is used to convey feedback there are concerns that there can be a trade-off between being able to immediately check for understandings, to engage in restatement and rephrasing of messages, and being able to capture a record of the communication outcomes for later reference.

Politeness and face management: it has already been noted that remote email communication lacks the paralinguistic elements of face-to-face communication and takes more time to construct than face-to-face communication (King, McGugan, & Bunyan, 2008; Münzer & Holmer, 2009; Stannard, 2008). This can have implications for the levels of between-

examiner familiarity as there is less information shared and it may also lack nuance. According to Cornelius & Boos (2003), familiarity influences how participants make sense of communication. It is possible that examiners who are familiar with each other will have accrued more information about each other and will understand each other's needs more than those who are new to each other. The issue of the literal interpretation of email messages (e.g. lacking the nuance of spoken discourse) and the inability of a feedback giver to check an immediate reaction to their feedback communication can be a challenge for face management.

The recipient having trust in the feedback giver: The hierarchic professional structure implicitly confers status to the feedback giver. This status is based on their experience and accrued expertise as an examiner. Whilst this status may translate as 'trust', it may be the case that the narrow bandwidth of remote communication may result in communication not addressing the needs of the examiner, and this could undermine their trust in the expertise of their TL (e.g. Liberman, Sagristano, & Trope, 2002). This issue also has implications for cases where a TL has not previously worked with an examiner. TL-examiner familiarity may also be an area where trust is implicated and where narrow communication bandwidth may adversely influence how a new TL and examiner partnership establish common ground with each other.

Positive information feedback: It has already been noted that hierarchic working arrangements mean that a feedback giver has a number of people to communicate with in an intense period of time, and that this can place a strain on communication. In this situation, a TL's feedback giving needs to be quick and concise so as not to interfere with their own process of marking examination scripts. These commitments make it likely that feedback messages will focus on what the TL considers to be the most salient issues for discussion, and these are likely to be the marking disagreements that are the spur for the feedback interaction. This time critical focusing may represent a hindrance to the sharing of positive information, or to a more dialogic interpretation of the TL role.

To summarise, in this chapter I have described the literature gathering and review methods that I used to construct my Integrated Analytical Framework (IAF). This framework outlines the factors that potentially influence the way that feedback supports the alignment of perspectives, which I define as an indicator of effectiveness. My framework overcomes the limitations of previous conceptualisations of feedback that are underpinned by behaviourist

approaches to communication and learning through the identification of ‘core’ (general, standalone) and ‘socio-cultural’ (context-bound) feedback factors. By outlining these factors, the IAF presents a picture of feedback-giving as a complex practice, and provides a resource for evaluating the effectiveness of feedback data.

In the next section I discuss my theoretical framework. I also outline the specific context of my study as a precursor to drawing connections between the IAF and broader psychological theory relating to communication and sociocultural learning theory.

3. My Theoretical Framework

In this chapter I describe how I develop my theoretical framework for my study. According to Silverman (2004), a theoretical framework provides 'a footing for considering the world, separate from, yet about, that world' (2004, p.14).

For my theoretical framework I move from the understandings that I gleaned through the literature review (described in Chapter 2), to outlining how I capture and make sense of my study data (examiner feedback messages) to answer my two research questions:

1. What is examiner feedback?

2. What is effective about such examiner feedback?

This move from literature review to data analysis involves a number of steps, which I describe in this chapter.

Following an outline of the examiner feedback context, which sets the scene for my study, I describe how I link the outcomes of the literature review to extant sociocultural learning theory, and, in particular, Mercer's (2002; 2008b) concept of the Intermental Development Zone (IDZ). This theory draws attention to how the nature of communication influences learning, and therefore the importance of the study of communication for studies of learning.

I then outline how this theory steers my methodological decisions when considering how to structure my data collection. In particular, I outline how a methodology that is allied to the sociocultural learning approach, termed Sociocultural Discourse Analysis (SCDA), focuses enquiry around four communication concepts: Content, Time, Joint Intellectual Action, and Impact.

Finally, I outline my approach to Sociocultural Discourse Analysis (SCDA), which I extend by drawing on a variety of methods that I unify under the umbrella term 'augmented Sociocultural Discourse Analysis' (ASCDA).

3.1 The Professional Examiner Context

I start this chapter by describing the professional examiner feedback context. This is important because it allows me to draw attention to a central problem that spurs my study. In describing the professional examiner feedback context I explain that Team Leader (TL)

feedback practice is a new and hidden practice. I also explain that feedback is expected to perform critical quality assurance (QA) functions that underpin public confidence in the assessment outcomes delivered by the Awarding Body. These concerns inspire my research questions, which are to establish what examiner feedback is, and what is 'effective' about it in terms of bringing together an alignment of understandings, so as to inform future examiner training and guidance.

I need to restate at this point, for the sake of clarity, that the definition of effective feedback that I adopt (i.e. feedback which contributes to the alignment of perspectives or the building or maintenance of common ground) is my own definition and is embedded in relevant literature that I draw on from beyond the context of my study of synchronous and asynchronous professional communication. This literature centres of how productive learning communication (i.e. that attains its learning purpose) relies on the development and maintenance of common ground (e.g. Beers, Boshuizen, Kirschner, & Gijselaers, 2007; Clark, 1996; Clark & Brennan, 1991; Edwards and Mercer, 1987; Mercer, 2000; Whittaker, 2003).

In England, Wales and Northern Ireland, educational qualifications are offered by designated Awarding Bodies. These organisations are recognised as being eligible to award qualifications by the national body that regulates qualifications and examinations, the Office of Qualifications and Examinations Regulation (Ofqual). The Oxford, Cambridge & RSA (OCR) Awarding Body is one of the main three awarding bodies in England, Wales and Northern Ireland.

The large scale delivery of school examinations in the UK is a highly regulated practice. The Code of Practice issued by the examination system's regulatory body in England (Ofqual, 2011) is an important influence on the ways that examiners interact. The Code stipules the quality assurance (QA) processes that awarding bodies need to adhere to for accreditation of their qualifications. Accreditation is important to awarding bodies as it unlocks access to an indirect stream of government revenue. Schools and other educational institutions can only claim funding from the government for those candidates who are entered for Ofqual accredited examinations. This funding is then passed on to awarding bodies as payment for delivering and processing the examination.

In contrast to models of assessment delivery where assessors are physically co-located, e.g. South Africa (Department of Basic Education, 2009) and Australia (Ofqual, 2013), examiners

in England generally mark examination scripts from their own workspace. Traditionally, boxes of examination scripts would be delivered to the examiner for marking and returned by post. Technological developments mean that these examination scripts are now delivered and marked as digital objects. Reflecting this distributed assessment model, the Ofqual Code's QA model centres on the standardisation of these physically distributed examiners' marking judgements. The purpose of standardisation is to 'ensure that all examiners have a well-founded and common understanding of the requirements of the mark scheme… and can apply them reliably' (Ofqual, 2011, p.25). The Code also bases the marking standard on hierarchic principles, e.g.

> *'[The] professional judgements [of the principal examiner] on the interpretation and application of the mark scheme for the unit/component must be final. Where there are large numbers of examiners and a supervisory structure has been set up, the awarding organisation must ensure that coordination of all assistant principal examiners and team leaders takes place to ensure consistency of practice' (Ofqual, 2011, p. 25).*

Like other large scale awarding bodies operating in England, OCR organises its examiners into hierarchically organised panels, with examiners' performances being supervised by a senior examiner (termed a team leader [TL] in OCR). The hierarchic element of the Ofqual Code is mirrored in the guidance given to examiners by OCR, i.e. 'OCR requires all responses to be marked exactly as the Person leading the marking would mark them' (OCR, 2012, p. 18). Figure 3.1 below outlines the marking hierarchy implicit to the panel structure, with the most senior examiner (the Principal Examiner) devolving examiner monitoring responsibilities to other senior examiners (TLs).

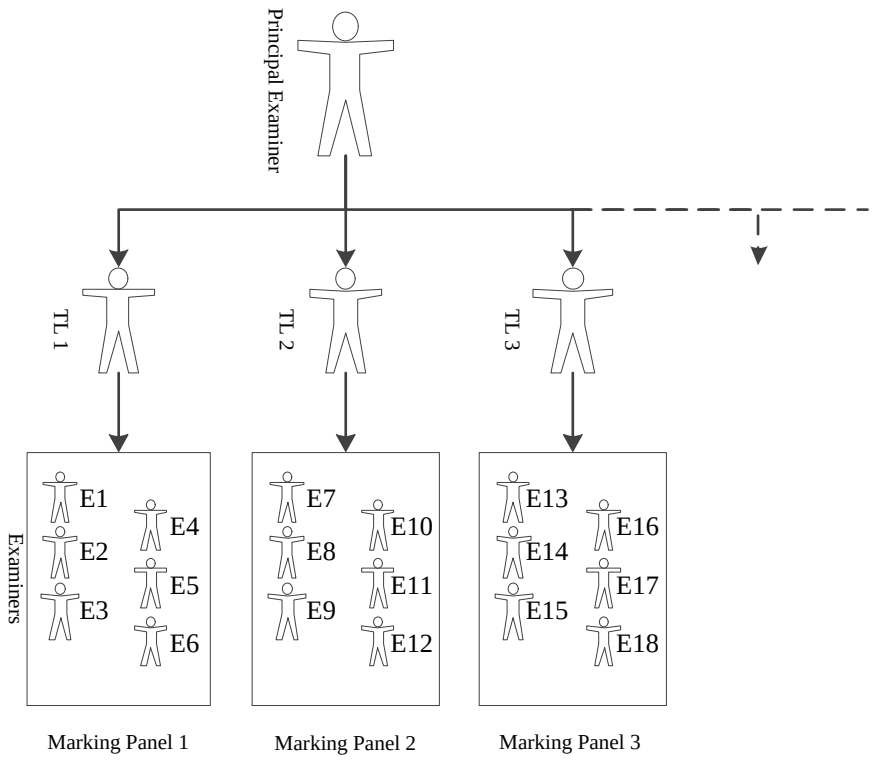

Figure 3.1: The OCR Marking Panel Structure

There is evidence from earlier research that between-examiner disagreement is a characteristic of marking systems that employ multiple examiners, and that this may be related to the need for examiner interpretation during the marking process. For example, research suggests that disagreement relates particularly to the level of complexity of some types of items (exam questions), with more objective items being subject to higher between-examiner agreement compared with other items (Massey & Raikes, 2006). Bramley (2008) goes on to argue that this disagreement relates to the potential level of constraint that pertains to the item, with less agreement being found where there is less constraint on the acceptable answers allowed by the item.

Research also suggests that standardisation has an important role in helping examiners' views converge around the features that constitute good or poor quality performances (e.g. Meadows & Billington, 2005). The cited research suggests that this effect relates to the opportunities that arise from interactions where more senior and less senior examiners discuss their understandings of mark schemes, and less senior examiners receive feedback on their mark scheme application, i.e. through focused discussion about examiners' particular marking decisions on specific student performance scripts.

Here, I investigate the examiner interactions that are at the centre of the OCR QA mechanisms. Seen in its entirety, the examination process involves examiners at different levels of seniority across a number of stages, from question paper setting through to the issuing of results. Figure 3.2 below shows an overview of this process, adapted from OCR (2013), locating the area of interest for this research (indicated by the box with the dashed line). This is the area where there is the greatest level of interaction between senior and other examiners. During this stage of the process, the QA arrangements focus on training examiners to apply a mark scheme that has been previously agreed by senior examiners. This stage also involves the senior examiners monitoring the marking performance of the examiners in their team, and giving the examiners feedback on their performance of applying the mark scheme.

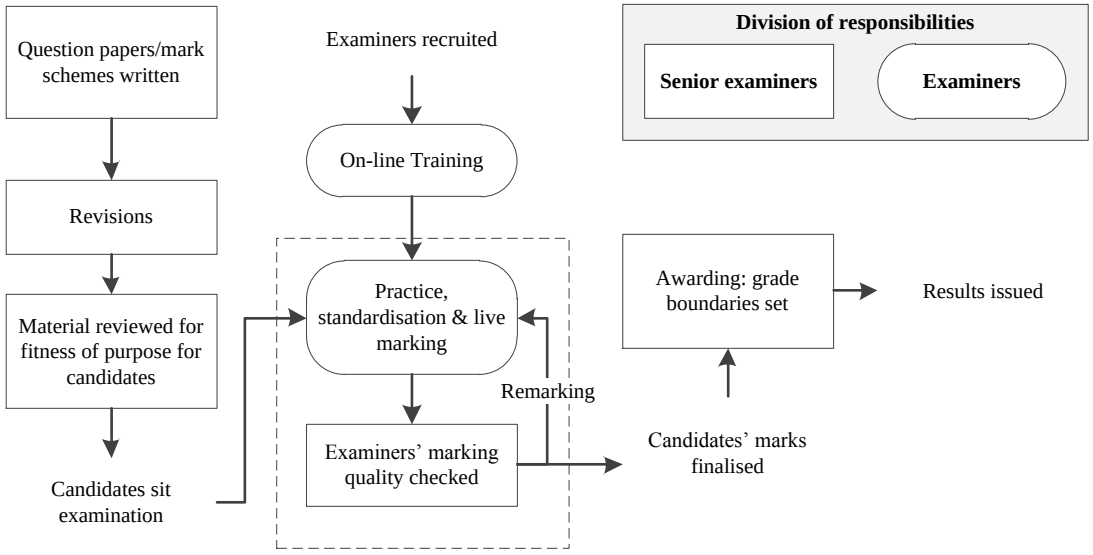

Figure 3.2: OCR Marking Process (Adapted from OCR, 2013)

3.1.1 The E-Marking Environment

OCR started to develop a digital marking system in 1999. This development was in tune with moves by the other large awarding bodies, who had also started to develop their own digital marking systems in the early 1990s (Ofqual, 2013). Cambridge Assessment, OCR's parent organisation, originally intended to purchase a custom built digital assessment delivery system, but early discussions with assessment agencies overseas failed to identify a system that suited the particular needs of the UK environs. As a result, Cambridge Assessment developed a custom-built digital marking system, *scoris® assessor*, in conjunction with the education technology development company *RM™ Education*.

The *scoris® assessor* system was piloted with a limited number of qualification units for live marking in 2006. Piloting then led to a larger roll out across more qualification units from 2008 onwards. The marking system has also now been employed in a range of other national contexts, with claims that it is…

> *'…the world's most widely used onscreen marking application. It is deployed by some of the most highly respected awarding organisations across the globe. In 2013 examiners in more than 76 countries used the system to e-mark 143 million exam pages' (RM Education, 2014).*

The *scoris® assessor* digital marking system affords a number of QA measures. The simultaneous distribution of digitally scanned versions of common examination scripts across different examiners allows examiners' marks to be compared with each other in ways that were not previously practical. Examiners' marking data is also available in real time to senior examiners and OCR administrators, who are jointly responsible for ensuring that high quality marking is maintained throughout the marking period. Another benefit of the digital marking system is that senior examiners can engage more frequently with the examiners in their marking panel by giving them feedback on their recently completed marking performance. These benefits are reflected in an Ofqual report on marking which states:

> *'As well as its logistical benefits, on-screen marking should improve marking reliability by enabling more frequent and flexible monitoring of examiners by exam boards. Senior examiners review their team's marking almost in real time, ensuring that inconsistent or inaccurate marking is detected early' (Ofqual, 2013, p.12).*

In the *scoris® assessor* marking environment, examiners and TLs remotely interact around practice, standardisation and live marking scripts (Figure 3.3). During these marking phases the examiners have access to a mix of generic and individualised feedback on their script marking from senior examiners. The aim of this feedback is to help examiners to learn how to consistently apply the mark scheme. The intensity of the feedback delivered to examiners alters as they move from practice and standardisation marking through to live marking. In the earliest marking stages feedback focuses on groups of 10 scripts. During live marking, feedback focuses on individual scripts that are randomly spread at a ratio of 1:20 throughout an examiner's marking allocation. In the case of the examiners in this study, a marking allocation would be around 200 student scripts. This difference in intensity reflects an assumption within the system design that earlier feedback aims to help an examiner to learn

how to apply the mark scheme prior to approval for live marking, whilst later feedback aims to remediate particular marking issues as they are encountered.

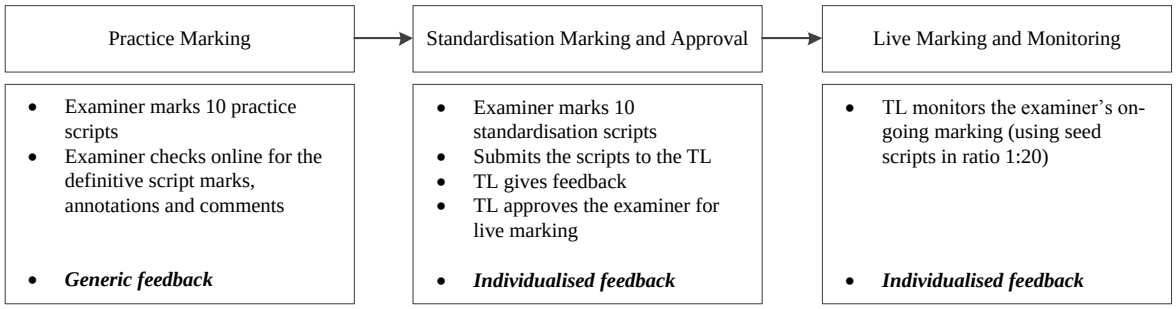

Figure 3.3: *scoris® assessor* Marking and Feedback Processes

Closer inspection of the marking and feedback stages shows the division of labour within the marking hierarchy. Figure 3.4 expands on the OCR Marking Process outlined in Figure 3.3. Senior examiners/TLs take part in a Standardisation Set Up (SSU) meeting where they agree on the final version of the unit's mark scheme, choose scripts that will be used for examiner practice, standardisation and monitoring, and fix an agreed definitive mark to these scripts. Other examiners gain insights into the appropriate way to apply the established mark scheme through marking the chosen SSU scripts and receiving feedback from their TL on their marking.

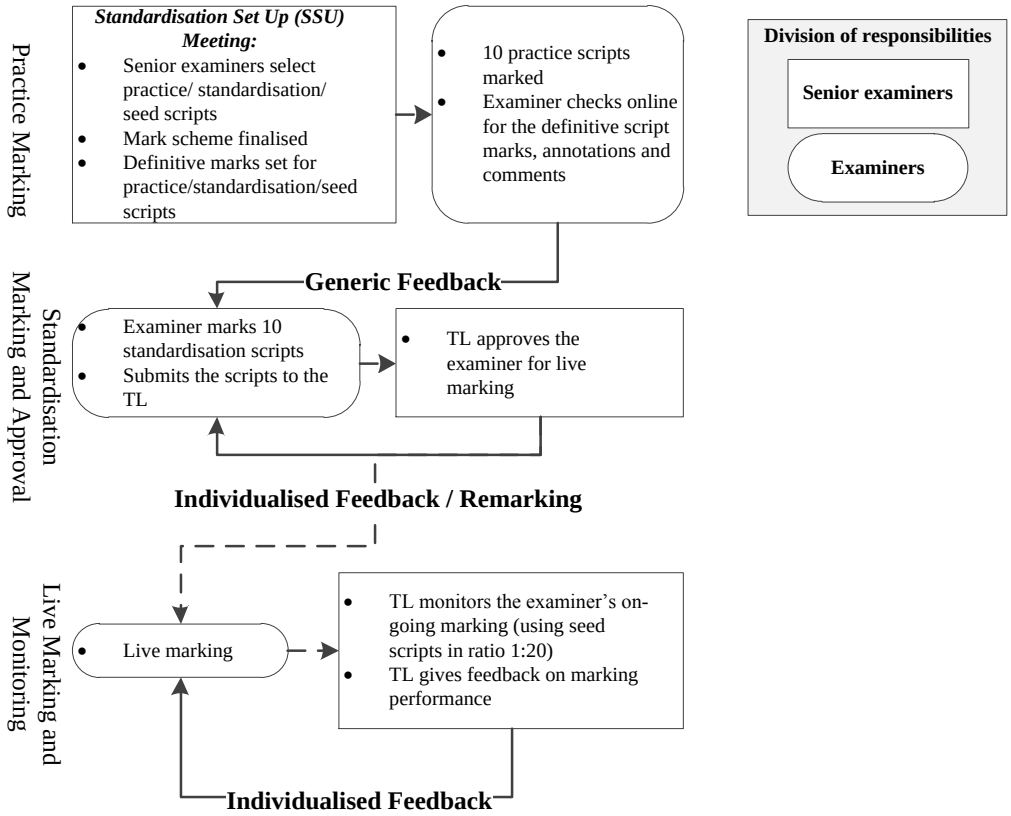

Figure 3.4: The Division of Examiner Labour across the Marking Phases

Figure 3.4 also shows how marking feedback is either generic or individualised. Feedback on practice scripts comprises a set of comments preloaded by the SSU meeting panel. An examiner needs to look at these comments to help them understand why their mark might differ from the definitive mark set by the SSU meeting. This means that the feedback is generic, as it is common to all examiners and does not link to the actual mark given by any individual examiner. On the other hand, feedback on standardisation and seed scripts is individualised, being given by a TL to an examiner in their marking panel, and relating to a particular mark given by the examiner. The communicative channels of the *scoris® assessor* system maintain the hierarchic marking structure. Feedback communication is restricted to linear pathways so that examiners can only send messages to their monitoring senior examiner (and not to other examiners at their own status within their marking panel). This means that the majority of interactions that take place throughout the system as a whole are likely to involve communication between senior examiners and examiners. In particular, these interactions are likely to involve feedback communication around practice, standardisation and live script marking.

I argue that the professional feedback context that I have outlined implicates a number of problems and that these spur the research questions that underpin my study. Examiner feedback is a new practice that has been under-researched. This lack of research reflects the sensitive nature of research into actual examiner practices, and its associated ethical issues (this is discussed in more detail in Chapter 4), as well as the information security issues that are tied into the development of examiner communication practices. An affordance of digital marking systems is that access to information can be restricted to those roles that require it, and thus potentially supporting efficient practices (through reducing information overload and superfluous communication). This paucity of research therefore makes it important to gather information about feedback so that it is no longer a hidden practice. This transparency then enables me to explore, through reference to theory, prior literature, and empirical participant data, what is effective about it and how to make it better.

The regulated environment in which the AB operates requires that QA indicators are used to monitor the functioning of the process so that it is fit for purpose (i.e. delivering fair and equitable assessment outcomes). One of these indicators is the level of between-examiner marking agreement, as this is a measure of marking reliability which itself is a component of validity. Feedback has a key role in supporting this alignment, and aims to ameliorate marking disagreements through sharing sanctioned hierarchic interpretations. One potentially problematic issue is that the assumption of mark stability that is the foundation for between-examiner alignments might be tenuous. Examiner disagreement can be linked to the complexity of the marking judgements required when applying a mark scheme (Bramley, 2008; Massey & Raikes, 2006), and level-based, holistic mark schemes may be particularly prone to disagreement as they include a degree of subjective examiner judgement. This complexity potentially implicates an ontological issue around the questions of whether a ‘true’ mark actually exists for a performance, and examiners might need to put this concern on hold in order to reach an agreement with other examiners.

The foundational content for communication is important to consider because, some would argue, reflexive discourse is a means through which participants deal with problematic ‘trouble sources’. According to Koschmann, Kuutti, & Hickman (1998) discourse involves self-righting mechanisms where any breakdown in understanding is a spur for communication repair, and an opportunity to ‘make visible aspects of the situation that might otherwise elude awareness’ (Koschmann et al., 1998, p.28). This visibility helps to assure that participants understand each other’s perspective as discourse unfolds (Schegloff, Jefferson, &

Sacks, 1977; Schegloff, 1992). Therefore, discourse, when seen from this perspective, is shaped by the existence of trouble sources, and, I argue, mark instability is a potential trouble source that feedback discourse seeks to repair.

The issue of disagreement around an unstable object implicates two problems that relate to the establishment of common ground through communication. Firstly, feedback can be considered to be a mechanism for professional learning (Basturkmen et al., 2014; Mathiesen, 2012). Feedback is a form of communication that implicates shared language and ways of thinking that helps to bring less experienced participants into a community of practice (see Chapter 1 for more on this). This learning process requires the building and maintenance of common ground between participants, and if this common ground is weak it is possible that communication will break down, the participants will fail to establish shared understandings, and the learner will not become a full participant in the professional community.

The structure of the examiner hierarchy means that there are aspects of information surrounding mark scheme generation and application that are exclusive to those senior examiners (i.e. TLs) who attend the SSU meetings. Feedback on examiner marking is a mechanism for reinforcing the practical application of the mark scheme, but, as I have demonstrated in the literature review in Chapter 2, the affordances of remote CMC can involve media-poor features that are a weak basis for attaining aligned understandings.

The second issue raised by the problem of communicating around an unstable object is that there is the potential for high levels of disagreement, as there are likely to be different perspectives prevailing around the object. For example, examiners may have different views on why a mark should be awarded if they have not had access to the same conversations around mark scheme application. In this sense, feedback is a form of discourse repair act that is fuelled by misalignment. This misalignment makes the existence of feedback communication a necessity because feedback is acting as a self-righting mechanism to overcome the sorts of disagreements that are a threat to assessment validity.

A corollary of the use of discourse to construct common ground around disagreement is that there are likely to be implications for relationship management. As I demonstrated in Chapter 2, a focus on negative information is not generally indicative of effective feedback, and this is exacerbated by remote communication technologies that fail to convey some of the paralinguistic cues that enable nuanced interpretations of messages. It would be anticipated from other literature (e.g. Goffman, 1967) that there may be managerial work involved in

maintaining professional face so as to enable on-going examiner engagement in learning discourse (and to ensure that examiners can develop to reach their potential through learning feedback). This is discussed further in the next section of this chapter, where I relate the literature reviewed in Chapter 2 with other theory.

In summary, I have outlined how TL feedback is a new and hidden practice and that it has dual QA and examiner learning functions. These considerations inspire my research questions, which are to establish what examiner feedback is; and what is 'effective' about it in terms of bringing together an alignment of understandings (and to inform future examiner training and guidance).

My description of the marking context implicates an ontological concern about mark stability, suggesting that this represents a trouble source that motivates participants to repair through feedback discourse. A concern for my study is that this trouble source is a weak foundation for building shared understandings. Moreover, findings from my literature review suggest that this weakness is exacerbated by structural issues (e.g. restrictions of access to shared information), technical issues (e.g. media-poor communication channels), and emotional issues (e.g. negative communication and face management).

In my theoretical framework social interaction is central to the study of language, and because of the above considerations the features of effective feedback that I have previously outlined require substantial further investigation. In particular, the idea of what constitutes 'Discourse that seeks to support the construction of learning (Pedagogy factor), rather than merely transmitting important information' needs to be explored in detail, since this seems crucial to all other elements of effective feedback, and to an 'ethos' of shared responsibility for agreement around marking.

3.2 Relating the Literature Review Outcomes to Theory

In this section I link the feedback features associated with effectiveness in the literature review to sociocultural learning theory in order to explain the basis for any such effectiveness. To do this I introduce Mercer's (2002; 2008b) Intermental Development Zone (IDZ), which is a model of learning communication that is situated in sociocultural learning theory. I also highlight the potentially problematic aspect of misalignment in examiner feedback and how this might implicate the need for additional theory from beyond a

sociocultural learning perspective in order to explain how feedback attains its productive outcomes.

Sociocultural perspectives suggest that communication and language use is a crucial area of study for understanding the development of individual thinking and learning that derives from social interaction. Although sociocultural research is still exploring how language influences the transformation of reasoning (Littleton & Mercer, 2013, p.99), the perspective suggests that language supports the development of participants' reasoning through the alignment of culturally appropriate collective thinking.

This perspective is inspired by the work of Vygotsky (1978), who proposed that social interactions lead to the development of participants' thinking, with ideas shared on the social plane coming to be internalised within the thought structures of individuals. Therefore, the development of thinking in an individual (perhaps conceptualised as learning or as the alignment of perspectives with another) is a product of the nature of the interaction that takes place between individuals.

Sociocultural perspectives are based on a dialectic model of how development occurs through interaction. This dialectic model, linking sociocultural thinking with its Hegelian and Marxian heritage, suggests that the development of ideas involves the confrontation of opposing states of thought. According to Ilyenkov (2008) the dialectic process involves,

> *'elucidating contradictions and of concretely resolving them in the corpus of a higher and more profound stage of rational understanding of the same object, on the way toward further investigation of the essence of the matter' (Ilyenkov, 2008, p. 190).*

In other words, misaligned thinking that, as a form of internal contradiction, encourages further interaction that seeks to resolve such misalignment can fuel social interactions. This idea has similarities to the concept of cognitive dissonance (Festinger, 1957) and coheres with observations that reflexive discourse involves self-righting mechanisms (e.g. see Koschmann, Kuutti, & Hickman (1998) mentioned earlier).

In this sense, examiner feedback discourse is constituted and refined through the interplay of participants' contrasting perspectives, with examiners having a shared imperative in reducing misalignment. In line with the work of Edwards and Mercer (1987), feedback, like other communication forms, is a site where participants endeavour to establish common ground

with each other so as to establish a basis of productive (i.e. on-going and purposeful) discourse. To do this they can draw on shared resources that invoke concepts that they believe to reside within the cognition of others and navigate around areas of privileged knowledge (i.e. knowledge that pertains only to one participant). For examiners these resources would be expected to include things such as subject knowledge (also known as background common knowledge) and references to any past interactions to create new dynamic common knowledge that is particular to the interacting participants.

According to Mercer (2000) and Littleton & Mercer (2013), who have looked at productive discourse in terms of talk that is effective for learning, there are some forms of talk that are better suited for learning compared with others. Disputational Talk is the least productive form of discourse as it is dominated by assertions and counter-assertions, disagreement and individualised decision-making. On the other hand, Cumulative and Exploratory Talk involves the participants making judicious linguistic choices that reference features of the shared context (Clark, 1992) and encourage shared, collective thinking in the ways that they allow 'ideas [to be] shared openly and possible explanations [to be] considered critically but in an atmosphere of trust' (Littleton & Mercer, 2013, p. 93).

To explain how productive social interaction attains individual (intramental) development Mercer (2002, 2008b) and Mercer & Littleton (2007) theorise that participants use shared resources to establish and maintain a shared framework of mutual understanding called an Intermental Development Zone (IDZ). In the IDZ, 'participants' minds are mutually attuned' (Mercer, 2002, p. 143). In a learning context the establishment and maintenance of the IDZ is the primary responsibility of more expert participants, and this has parallels with professional examining contexts where the TL has responsibility for ensuring that other examiners are attuned to their thinking.

Mercer (2002) observes that talk is the principal tool for creating the IDZ, although I would also argue that other forms of discourse using cultural tools (e.g. email interaction) have the potential to support this creative process. This is because email interaction affords an iterative ease (initiation and response can be carried with a single click of a button), and it is based around words which have semiotic qualities that encourage participants to share a focus and construct meanings. Wertsch & Kazak (2011) highlight how words are important mediating tools at the heart of learning because they have the potential to have double readings. This

affordance means that words are able to engage simultaneously with concepts at different levels of sophistication:

> *'[words] are incredibly robust in that they can allow interpretation and understanding at many different levels yet still support the intermental functioning required to move learning and instruction along' (Wertsch & Kazak, 2011, p. 156).*

This means that both experts and novices (or in terms of the community of practice metaphor, those at the centre and those at the periphery of a community), can simultaneously interact around the same object of text, and start to create a shared framework of understanding. In this way, texts act as boundary objects that allow community members to construct and refine their shared perspective (Boland & Tenkasi, 1995). Although these forms relate to spoken discourse, it is possible that they also have relevance to the spoken and written forms of examiner feedback discourse.

At this stage I would like to look beyond conventional sociocultural theory to consider how these apparently incongruous forms of discourse may be managed. I have used the concept of 'articulation', borrowed from Strauss (1985) and adapted by Schmidt (1994, 2011), to simultaneously reference the notion of 'expressing' and the act of 'coordinating interconnected work across individuals', and argue that feedback can be used to perform both functions (Johnson, 2015). According to Strauss and Schmidt, articulation work involves the often unnoticed and taken for granted work that is carried out by managers to ensure that those around them complete their own tasks, and thus ensure that mutually important strategic goals are attained.

In the examination marking context it is likely that TLs will be very conscious of the need to ensure that competent examiners (or those who the TL believes will become so) are motivated to complete their marking tasks. The consequence of losing competent examiners from the workforce before marking is completed represents additional workload for the TL, and a concomitant strain on their relations with other examiners who then share the additional workload.

In the context of misaligned and negative discourse it is possible that TLs need to attend to the macro-function of language to attain the effective management of relationships (Brown &

Yule, 1983) and to the interpersonal skills that allow TLs to 'develop bonds' (c.f. Carpenter & Wisecarver, 2004) with relevant examiners in order to ensure job completion.

In summary, based on the sociocultural learning theories that I have drawn on I see it as legitimate to claim that feedback is a collaborative process and is not an unproblematic transmission of information. This process is a dynamic collaborative intellectual activity where participants in discourse use background common knowledge as a resource to create productive (learning) interaction.

From this sociocultural perspective, which coheres with the messages that I have gleaned from my literature review when creating my Interpretive Analytical Framework (IAF), it is absolutely fundamental to consider language interactions and how they establish a basis for the optimal operation of the features that contribute to effective feedback practices. In this way my study considers, in depth, the meaning of the words 'discourse that supports the construction of learning' (rather than the mere transmission of information) in this context. This feedback is participative, with an active recipient role, and assumes that the learner engages in self-regulation in the learning process. I also noted in the previous chapter that effective feedback tends to be immediate and use clear and specific language. Finally I note from the literature that these characteristics tend to be easier to achieve in collocated interaction situations that employ rich media communication channels.

In highlighting the two-way communication process that characterises effective feedback my IAF also highlights that feedback contains some implicit elements that centre on relationship management alongside the explicit elements that seek to codify important information. It also draws attention to the emotional issues that pertain to the communication of negative information and how the use of politeness in face management appears to be an element of effective discourse.

As a consequence, a central hypothesis driving the empirical research undertaken here is that IAF factors that support the development of an IDZ enhance feedback effectiveness, whereas factors that undermine the development of an IDZ lead to ineffective feedback. The establishment of common ground underpins the development of an IDZ, and below I outline how the features identified in the IAF (Chapter 2) relate to the construction or undermining of common ground.

3.3 Moving from Theory to Method

Having outlined the processes through which social interaction influences individual development, I now look at the requirements of a sociocultural discourse analysis (SCDA), which Littleton & Mercer (2013) argue involves three perspectives: linguistic, psychological and cultural. This in turn leads to consideration of the use of a mixed methods research approach.

From the linguistic perspective, data needs to be collected on the content and qualities of the participants' discourse in a representative sample of cases. My interpretation of SCDA suggests that these qualities need to be considered in four areas.

- Content - the lexical content and the cohesive structures of discourse;
- Time - how shared understanding is developed in social context over time;
- Joint Intellectual Action - evidence of how participants recognise the thinking of other participants and use this to coordinate shared understandings;
- Impact - the effects that discourse has on the cognition and behaviour of the participants, which in the case of examiner feedback is to attain a point of agreement or resolution.

The first area that SCDA needs to consider is the content base that participants in interaction use as a shared resource for common ground building. The second area that SCDA needs to evidence is how discourse changes over time so as to gain insight into the way that common ground builds and develops. According to Mercer (2004), this is an overlooked area since 'few researchers have tried to relate the content, quality, and temporal nature of discourse in joint activities to outcomes' (Mercer, 2004, p. 139).

The third area of interest is to gather evidence about Joint Intellectual Action and how it gives insights into how participants attend to the needs of other's in their own discourse. The final area that SCDA needs to consider is the direct impact of the interaction on outcomes (e.g. the alignment of individual perspectives).

SCDA not only defines the areas of interaction that are of interest for a sociocultural analysis, but it also offers a blueprint for the approach needed for data collection. According to Littleton & Mercer (2013), the analysis of content and qualities over time requires the use of both qualitative and quantitative methods. Such an integrated approach allows information to be gathered about the particularities of discourse whilst also allowing these to be related to larger chunks of data to support broader generalisation. In this way, quantitative data 'is taken to aid the understanding of the qualitative, as opposed to the converse' (Knight & Littleton, 2013, p. 2). Sociocultural perspectives legitimise the use of mixed qualitative and quantitative

methods in research design. For example, whilst descriptive statistics can allow insights into demi-regularities between phenomena, it is also recognised that qualitative methods can give insights into influence because they tap into the context-dependence of social phenomena (Zachariadis, Scott, & Barrett, 2010).

Whilst linguistic analysis of the content and quality of participants' lexical and cohesive structures is an important element of a sociocultural analysis, analysis also needs to consider the data from a psychological perspective. This is because communication has both enacted and perceived dimensions, with the development of thought involving language at inter- and intramental levels (Vygotsky, 1978). This inter-relationship means that the researcher needs to consider how language impacts on participants and how they construct meaning in discourse (e.g. looking at rationalisation strategies or how language is used to control thinking within interaction).

Mercer (2004, 2008) also privileges the importance of evidencing the outcome of interaction in sociocultural analysis. This evidence helps to ensure that an analysis moves beyond a description of how language functions towards an explanation of how discourse may impact on cognitive development. In my approach to analysis I term this element 'impact', and this concept has implications for research methods. The measurement of impact on participants' thinking this might be evidenced in both qualitative and quantitative terms. For example, the impact of discourse may be explored through statistical tests of performance outcomes (Mercer et al., 1999), or through participants' perceptions of communication (e.g. Johnson & Black, 2012).

Finally, the data needs to be considered from a cultural perspective, so that any contextual features that may influence meanings in the discourse are taken into account (e.g. any standard operating procedures used within a professional work context).

Building on the theory outlined in the previous section, the areas of interaction that I consider to be central to SCDA can give insight into how participants establish and maintain common ground with each other during interaction, with common ground being the key mechanism for establishing an IDZ.

Based on the broad SCDA frame provided by Mercer (2004) and Littleton & Mercer (2013) I now look more closely at the methods that I use to explore interaction. Reflecting the particular context of my study I chose to refine the mixed methods approach to SCDA

approach reported in earlier, largely classroom-based talk studies (e.g. Mercer, 2008). My study necessarily focuses on both written and spoken interaction, so I needed to be able to analyse both of these forms of communication as texts. My literature review evidence also made me conscious of the need to be able to consider the dynamic of negative interaction and how this implicates professional face

As a result I chose to adopt a methodology that incorporates a variety of specific perspectives to explore the four sociocultural areas of the original SCDA approach. I call my methodological refinement 'augmented SCDA' (ASCDA) (Figure 3.5). In the figure I conceptualise how the areas of interest to SCDA may be evidenced through the use of three approaches for analysing discourse: Conversation Analysis (CA), Thematic Content Analysis (TCA), which is a strand of Discourse Analysis (DA), and Corpus Linguistics (CL). This representation forms the basis of my ASCDA approach.

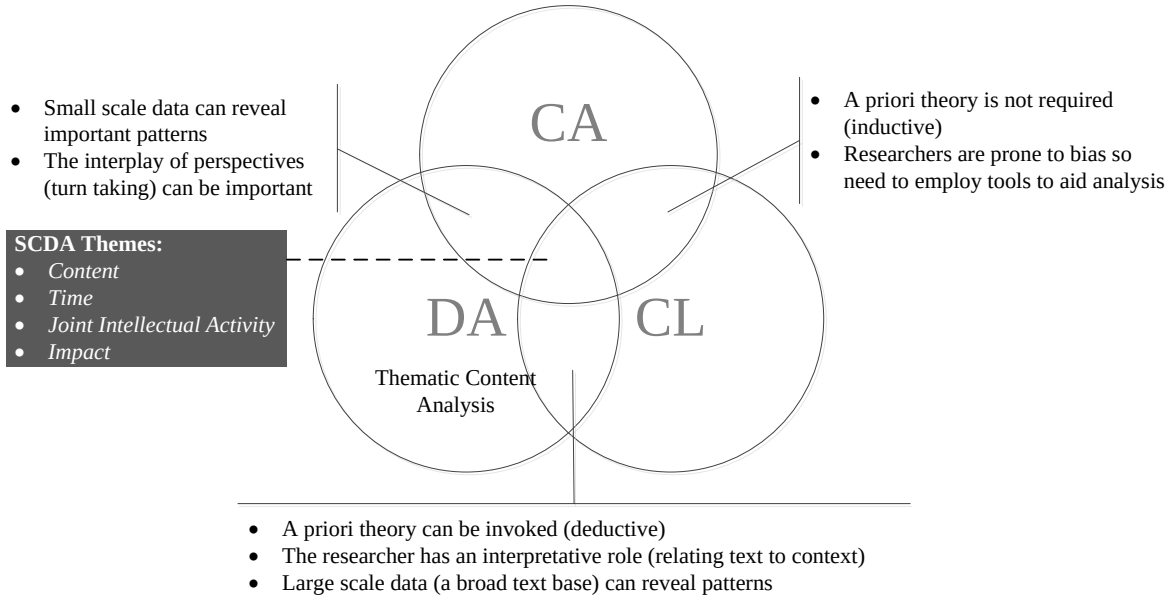

Figure 3.5: Areas of Methodological Overlap between the Approaches and the SCDA Themes

I acknowledge the philosophical and methodological challenges of bringing together methods from different methodological traditions (Johnson, 2016d), but also recognise the strengths of bringing together methods that can cross validate each other. Because I am interested in studying broad interactional elements of feedback discourse I look to the complementarities between methods that share a common focus on analysing how participants pursue social action goals through discourse.

In brief, CA-influenced approaches formed the organising frame for my data gathering and analysis because of their links to the way that they seek to evidence how relationships are implicated in the minutia of interaction data. This approach suited my concern about how my participants may deal with relationship management in the context of negative communication patterns. CA-influenced approaches are appropriate for analysing the detail of communication (e.g. content at word and phrase level), as well as the strategies that are employed by people in interaction (e.g. the ways that communication is structured to attain a desired effect). These strategies can be an indication of Joint Intellectual Action as they reflect the way that participants choose their communication content and structure to suit the perceived needs of the recipient, suggesting that they have an idea of the recipient in their mind when they draft the communication. CA-influenced approaches can also give insight into the direct impact of the interaction 'as experienced' by the participants, and this builds on the inductive approach to analysis that informs a CA methodology. Although it is an approach that is largely tied to the analysis of spoken interaction, I suggest that evidence of impact would also be evident when looking carefully at the handover points between email messages (i.e. when the onus to communicate is passed from one participant to the other).

TCA is an approach that can construct a broad picture of large-scale interaction data through rigorous and systematic coding. These codes can summarise the content of the discourse (at phrase level) and can be explored for indications of any inter-relationships across the whole discourse data set through quantitative or qualitative analysis. The coding process implicates the analyst in interpretation of the intentions behind any communication patterns, which also implicates the concept of Joint Intellectual Action.

CL –influenced approaches deal with discourse data at word-level and analysis looks to make apparent any patterns across a large body of discourse data. The use of specialist software allows the analyst to see numbers and patterns of word use in ways that are difficult to achieve through manual analytical approaches. The focus of CL is to allow an inductive approach to data gathering, with 'significantly used' words being made salient. With the use of further analyst interpretation these words can give an indication of communication content. Word search also allows the comparison of word use over time.

3.4. Summary of My Theoretical Framework

In this chapter I have described how I developed a theoretical framework for my study that draws on a sociocultural learning perspective. At a basic level, this framework seeks to

explain how feedback is a discourse mechanism that is expected to fulfil QA functions and reduce examiner disagreement.

A defining element of my framework is that communication study is a crucial area for understanding how individual thinking and learning derives from social interaction. In addition, the sociocultural learning theories that I draw on suggest that feedback is a collaborative process, and is not an unproblematic transmission of information. In constructing my framework I look beyond conventional sociocultural theory to consider how discourse is organised. To do this I use the concept of 'articulation' (Strauss, 1985) which describes how communication helps to coordinate interconnected work across individuals.

My theoretical framework suggests that feedback is made necessary by the existence of trouble sources, and that it seeks to resolve misalignment through the establishment of common ground between participants. These trouble sources are anticipated to include mark instability (e.g. highly subjective items items), a lack of shared information (due to media-poor communication modes), and a lack of shared experience (e.g. new and unfamiliar examiners). Finally, the framework suggests that feedback may also perform a relationship management function because of the potential interactions between the communication of negative information and the need to maintain 'professional face'.

The articulation work of building and maintaining common ground allows the examiners to create an Intermental Development Zone (IDZ) where their minds are mutually attuned. To evidence this I have augmented the SCDA approach to discourse analysis formulated by Mercer (2004) and Littleton & Mercer (2013). My approach, called ASCDA analyses discourse according to four themes of interest: *Content; Time; Joint Intellectual Action*; and, *Impact*.

4. Method

This chapter outlines the methodological framework that I adopt for my study. I start with a description of my pilot study and how this influenced the methods adopted for the main study. I then go on to describe the methods that I use in my main study, structuring this discussion around two research questions. For each research question, I outline the methods that I employed to collect and analyse the data, along with a rationale for the choices that I made when planning the study.

4.1 Pilot Study

I carried out a two-phase pilot study in January 2013. This pilot study investigated the features of marking performance feedback communicated between a Team Leader [TL] and a team of examiners. The pilot study had two questions:

- What does feedback comprise of?
- Are there variances in the amount of feedback, and why?

4.1.1 Pilot Study Method

In the first pilot phase I observed a face-to-face standardisation meeting to capture the nature of the interactions that took place during the meeting. The meeting involved 36 participants; seven TLs and 29 AS level[7] Economics examiners. The aim of the meeting was for all of the participants to reach common understandings about how to apply the mark scheme for the paper.

There was no opportunity to audio record the examiners' interactions, so I developed an observation schedule to collect data about the length of time that examiners were observed discussing particular items (Appendix C). This observation focused on one table of four participants; one TL and three examiners. The observation schedule also included space to note down instances of talk.

In the second pilot phase I observed three episodes where the TL gave remote feedback on their marking to each of the examiners from the earlier face-to-face meeting. These data

[7] A General Certificate of Education (GCE) A-level consists of four (or six for natural sciences) modules studied over two years. Normally, two modules are assessed in the first year, and make up a stand-alone Advanced Subsidiary (AS) level qualification.

comprised of two telephone feedback transcripts and one email feedback transcript. These transcripts formed a corpus of 8828 words.

4.1.2 Pilot Study Analysis

I used two types of analysis to explore the feedback data. My first analytic approach involved a Thematic Content Analysis (TCA) approach (for more on this see section 4.2.2). This method used MAXqda qualitative data analysis software (VERBI Software – Consult – Sozialforschung GmbH, 2013) to generate generic codes that captured the feedback characteristics. My coding framework was structured so that I could capture a number of discourse characteristics, namely:

- The content of each phrase (i.e. what information was conveyed);
- The referencing strategies used by participants;
- The questions used;
- Any tangible outcomes that were signalled in the discourse.

For the purposes of this coding process I used the level of utterance and word as the unit of analysis, as these could indicate moves in the discourse. Analysis of moves is a top-down approach that focuses on meaning and ideas, with each move representing a stretch of discourse that serves a communicative or semantic function (Upton & Cohen, 2009). I then explored the existence of structural regularities within the data at the level of 'exchange' or 'turn' (a practice that is usual within Conversation Analysis techniques, e.g. Gibson, Webb, & Vom Lehn, 2011). This approach was adopted to allow analysis to consider the texts at multiple levels; both top-down and bottom-up.

Appendix D outlines the Pilot Coding System that was developed to analyse the three episodes of feedback data. Forty-two sub-codes were generated during this analysis. These sub-codes were organised into six primary code categories that helped to group the sub-codes according to their purpose and position in the analysed discourse. My approach to coding allowed double coding of text segments to account for the multiple purposes that a singular text segment might be achieving. This is because a word that was highlighted by a TL in a feedback message could be interpreted as performing dual functions. For example, the underlined word in the following phrase '`Can you please review this…`' would be coded as both 'Accentuation' and 'Distancing' in my coding frame, since it acts to accentuate something of note as well as to convey something about the authority that is implicit in the relationship that the participants are constructing through their discourse.

My second analytic approach involved searching the data for particular terms that might be associated with 'Exploratory Talk' (Mercer & Wegerif, 1999). This analysis involved the use of Monoconc MP 2.2 concordancing and corpus analysis software (Barlow, 2012). The terms that were used to search the data as potential signifiers of exploratory talk were taken from Wegerif (1997), who notes that the hypothetical mode, which is 'essential to exploratory talk is usually served by conditional tenses introduced by a limited set of words' (Wegerif, 1997, p. 102). Wegerif points out that these words include 'if', 'might', 'would', 'could', 'should' and 'may'. Words associated with justificatory statements were also used as search terms, as it would be anticipated that feedback would include some justification for marking decisions. These justificatory words were 'because' and 'so', along with 'then' and 'but', which could also support the articulation of justification. Analysis of the corpus only included content words, and discounted (in)definite articles, grammatical contractions, question item index markers, and mark scheme abbreviations. The Stop List[8] that outlines these excluded terms is shown in Appendix E.

4.1.3 What Does Feedback Include?

Analysis of the TL talk types, although necessarily tentative due to the lack of audio recording, suggested that some discourse characteristics were used by TLs to help them to build shared meanings with examiners, albeit in a hierarchically driven way. The discourse characteristics that were observed included talk which:

- Offered new information;
- Referenced past experience;
- Requested information;
- Checked on the validity of information interpretation.

My analyses also suggested that there were structural regularities in the feedback discourse, with some codes commonly co- or near-occurring:

- There appeared to be a relationship between locating information for shared discussion, referencing back to previously referenced information (e.g. a mark scheme), and bridging the concepts in these information fields through exemplification;
- TLs and examiners appeared to share perspectives in connected moves. It is possible that the co-occurrence of TL and examiner perspectives represented a coming together and

[8] 'Stop Lists' remove specified words from being included in the analysis software. These are often grammatical or function words that are both of high frequency and generally not of interest for analytical purposes. According to Barlow (2012, p. 42) 'it is often desirable to omit them and concentrate on words of interest which might otherwise be masked by the more frequent forms'.

exploration of different perspectives. In some cases, there was a hint of the TL assuming prior knowledge of the examiner's perspective as part of their own thinking (e.g. 'that's probably because you thought…'). This suggests that inferred knowledge of the thinking of others could be part of an internal dialogic mechanism (Bakhtin, 1981) that might underpin some TL feedback processes;

- Giving information often led to the communication of additional information. This meant that an interjection into the discourse by a third party could spur further elaboration around an original point of information;
- Examiner acknowledgement of a point raised by a TL appeared to encourage on-going dialogue and TL flow through recognising the points being made and inviting the TL to continue with the information sharing process;
- Disagreement formed a specific focus for discussion. It was also interesting to note that there was also effort made by the TL to moderate the importance of any noted differences.

The identification of keywords and their collocated partner words suggested a number of functions for shared meaning building were being supported; examples included:

- Verification (e.g. '*So* your analysis is fine…);
- Clarification and emphasis (e.g. I agree with you *but* there are …);
- Rule transmission by TL (e.g. '*So* when / *If* you see the word "fixed" *then* you *should*…');
- Conjecture (e.g. '*If* one of these scripts turns up then you can…');
- Alleviating forcefulness (e.g. 'I *think* this mark is a bit out);
- Identifying creditworthy features (e.g. 'give a mark *for* / I *would* give a mark for');
- Explanation (e.g. '*so* give credit for / give credit *because* / I *think* this / this is *why* / we *thought* this was…').

4.1.4 Are there Variances in the Amount of Feedback, and Why?

There was evidence that discourse was item driven, with some items accounting for more discourse than others. In addition, particularly problematic items appeared to attract higher levels of discussion across all examiners. For example, most examiner time was spent discussing Item 6 and Item 3. These items had a relatively large mark allocation in relation to the total number of marks available on the exam paper, and one item used a levels-based marking scheme[9]. Taken together, these two factors might be considered as indicators of complexity that could lead them to becoming perceived as being problematic items.

[9] Levels-based mark schemes are those where the total mark awarded is determined by matching the quality, content, and combination of elements of the candidate responses to the appropriate level of response, each of which is associated with a band of one or more marks. Markers have to apply a principle of best fit. These mark schemes are considered to be more complex to use than objective or points-based mark schemes (Black, Curcin, & Dhawan, 2010).

The relative balance of the use of words with an exploratory potential in the corpus also indicated that the TL had a dominant role in the discourse. My data showed that the TL used these words more often than examiners, but that these words were used in similar proportions across the two sub-corpora (Table 4.1).

Table 4.1: TL and Examiner Exploratory Word Use

	TL corpus (84697 words)		**Examiner corpus (10789 words)**	
	N	% of TL Discourse	N	% of Examiner Discourse
For	1407	1.66	142	1.32
So	781	0.92	82	0.76
But	628	0.74	88	0.82
If	375	0.44	48	0.45
Because	211	0.25	19	0.18
Would	162	0.19	25	0.23
Think	137	0.16	30	0.28
Thought	59	0.07	14	0.13

These data suggest that a function of the feedback communication was to allow the TLs and examiners to outline their thinking to each other, with the TL accounting for most of this discourse.

4.1.5 Revisions Made as a Result of the Pilot Study

The pilot data gathering exercise informed the development of my main research study in a number of ways. At a general level the pilot study was useful because it demonstrated that a mixed methods approach could enable the elicitation of some of the particular and general features of examiner feedback interactions. The combination of coding and concordance analyses complemented each other to support analyses that allowed patterns to be observed across the whole data set, whilst also allowing space for my interpretive engagement with the discourse in some detail. Importantly, the coding framework generated in the pilot study provided a foundation upon which to build a refined coding framework for use with a larger data set in the next study phases.

Reflection on the pilot coding framework as I moved into the main study led me to introduce a number of changes to the coding framework that I used for my final analysis. My original pilot coding framework had six superordinate and 42 subordinate codes (Appendix D), and I refined this to 4 superordinate and 30 subordinate codes in my final coding schedule (Appendix H).

This refinement was partly a result of the greater amount of data that I had at my disposal in the main study, with my coding frame 'evolving' (Frankfort-Nachmias & Nachmias, 1996, p. 338) as I became more confident about the distinctions between observed elements of the feedback discourse data. My coding framework refinement also reflected the way that my understanding of sociocultural learning theory developed during the early stages of my research study, with a clearer link between SCDA and my coding structure reflecting the more focused thematic framework that underpinned my analysis.

Table 4.2 relates my original superordinate pilot code structure to the refined code structure that I generated for the main study. The table shows that there was more of an emphasis in the pilot study on codes that dealt with the SCDA *'Content'* concept, with some limited reference to the *'Impact'* concept. To broaden out the scope of the analysis I was able to incorporate the *'Structure'* and *'Style'* elements of the pilot coding framework into an *'Interpersonal Focus'* code, as I felt that these elements both addressed the way that TLs constructed their messages to create a desired effect on the recipient. Whilst being an element of content, and relating to the transaction of information, the *Interpersonal Focus* code allowed me to also account for the way that content was being conveyed, and therefore implicating the interactional dimension of language (Brown & Yule, 1983). I collapsed *'Disagreement'* into a more general *'Giving Information'* code, as I came to recognise that there was no theoretical reason to distinguish between this and other types of information (e.g. *'Giving technical information'*). This was because disagreement as a phenomenon was always necessarily linked to a piece of information, and needed this link in order for its interpretation to be made possible.

I also unpacked the original *'Information'* code into two new codes; *'Giving Information'* (which had a clear link to content), and *'Bridging'*. This second code described how TLs made links between information sources in response to the needs of examiners, which I interpreted as a form of Joint Intellectual Action, and as an indicator of movement through time as references to sources could be historical in character. Finally, I came to interpret the *'Verification'* code as a form of *'Action'* (or more specifically as an indicator of resolution) because I interpreted this as being an outcome of the confrontation of joint perspectives.

Table 4.2: The Relationship of Pilot Codes and Main Study Codes

Original (Pilot) Superordinate Code	***Sociocultural (SCDA) Reference***	**Refined (Main Study) Superordinate Code**	***Sociocultural (SCDA) Reference***
Structure	*Content*	Focus interpersonal	*Content*
Style	*Content*	Focus interpersonal	*Content*
Disagreement	*Content*	Giving information	*Content*
Information	*Content*	Giving Information/ Bridging Information	*Content* *Time/Joint Intellectual Action*
Verification	*Content*	Action	*Impact*
Post feedback action	*Impact*	Action	*Impact*

The changes that I made to my pilot coding framework reflect the way that this original framework was generated before I had fully developed my interpretation of SCDA; and the pilot framework lacked some focus on elements that I felt would be of importance (i.e. *'Time'* and *'Joint Intellectual Action'*).

Another refinement involved my decision to integrate a Stimulated Recall (SR) Interview phase into my feedback analysis. This addition reflects the sociocultural perspective that I adopt for this study (outlined in Chapter 3). Based on this perspective I claim that examiner feedback discourse is constituted and refined through the interplay of participants' contrasting perspectives, and that examiners have a shared imperative in reducing misalignment. Building on this, my interpretation of SCDA also steers enquiry towards evidencing 'Joint Intellectual Action' through gathering evidence of how participants recognise the thinking of other participants and use this to coordinate shared understandings. As a result, I felt that it was imperative that I should involve the perspectives of TLs and examiners when trying to understand what was effective about feedback discourse. This reflective interview strand drew on my insights gained from a previous study (Johnson and Black, 2012b), and I discuss the strengths and limitations of this method in a later section (4.3.1 'Problematising Interview and SR').

Finally, my pilot study gave me insights into the variances within feedback, which were helpful as a cue for focusing the area of enquiry in my main study. These variances appeared to relate to the complexity of item characteristics. Since disagreement was found to be a key element of feedback discourse, it would be anticipated that misalignment (and therefore feedback) would occur more around complex items. This insight is important because it

suggests that the focus for the main study should be on examiner feedback on examination papers which include complex question types. A study by Black, Curcin, & Dhawan (2010) discusses the characteristics associated with item complexity, and I used these characteristics to select the context for my main study.

Since disagreement appeared to be a key feature of feedback, I was able to reflect on how the context of misaligned and negative discourse may make it necessary for TLs to attend to (1) the macro-function of language to attain the effective management of relationships (Brown & Yule, 1983), and (2) to the interpersonal skills that allow TLs to 'develop bonds' (c.f. Carpenter & Wisecarver, 2004) with relevant examiners in order to ensure job completion. Again, as I have outlined earlier, this concern led me to explore how CA-influenced approaches could frame my data gathering and analysis because of their links to the way that they seek to evidence how relationships are implicated in the minutia of interaction data. CA-influenced approaches are appropriate for analysing the detail of communication (e.g. content at word and phrase level), as well as the strategies that are employed by people in interaction (e.g. the ways that communication is structured to attain a desired effect), and this allowed me to focus my attention at the handover points between email messages (i.e. when the onus to communicate is passed from one participant to the other) or in telephone communication.

I now go on to consider methodological issues associated with the two research questions that guided my main study.

4.2 Main Study Research Question 1: What are the Characteristics of Examiner Feedback?

The sociocultural theory that underpins my theoretical approach claims that knowledge is socially and collectively created by participants through communication. This creative process involves the participants using cultural tools. Vygotsky (1978) identified language as an important cultural tool; therefore, to understand knowledge construction entails the study of how communication and interaction work, and this has implications for methodology. One such implication is that the context of an interaction is a mental construct of the participants; something that they construct and reinforce and reference within their interaction. This links with the reference I made in Chapter 3 to the work of Gee & Green (1998), who argue that communication episodes influence and are influenced by their history and context.

This means that the methods used to study interaction need to be able to take into consideration the context that develops through and which influences that interaction. For example, an important element of context-in-interaction is the way that the instigation and reception of communication is intertwined, such that during the process of message generation a person is likely to have in mind a sense of the potential recipient reaction. Again, this accords with the work of Gee & Green (1998), who argue that meaning making is unavoidably situated, both physically and temporally, with individuals assembling images or conjuring up patterns from communicated messages on the spot and in light of their past experiences.

Therefore, evidencing how communication works requires the use of appropriate methods that can gain insights into the way that context is referenced within, and is an influence upon, communication. In this section, and the following sections of this chapter, I outline the methods that I use to ensure that I recognise the influence of context in communication.

The linguistic components of these interactions can be gathered and analysed as sequences of utterances. Yet exchanges between participants within a situated experience is the process of participants actively constructing inter-subjective meaning, wherein meaning making is 'an interactional achievement' (Stahl, Koschmann, & Suthers, 2006, p. 8). This means that participants' own subjective interpretation plays a key role in how communication works.

To capture the complexity of social interaction and communication, research methods need to be able to account for both its observable and its interpretative elements. As a consequence, and guided by my pilot study, I adopted a mixed methods approach for data capture and analysis. This approach is in line with a Sociocultural Discourse Analysis (SCDA) approach, which I outline below. For my analysis I further refined the concept of SCDA as initially developed by Mercer (2004) and Littleton & Mercer (2013) by integrating methods commonly associated with Discourse Analysis (DA), Conversation Analysis (CA) and Corpus Linguistics (CL) approaches. I chose to draw on these elements because they each contain features that are of interest to my study that derive from my sociocultural perspective. I have already indicated that I call my approach 'augmented SCDA' (ASCDA).

I decided to gather feedback data over a two-year period, which meant that data collection was carried out in two phases. I chose to focus on this extended period so that I could feel more secure that my analyses would capture stable characteristics of feedback (i.e. not be biased by any particularly anomalous features pertaining to a particular marking session in

any one year). This would make my analyses more robust and reinforce any claims that I might make about generalizability.

Phase 1 feedback data gathering took place between May and July 2014, and Phase 2 data gathering took place between May and July 2015 (Figure 4.1).

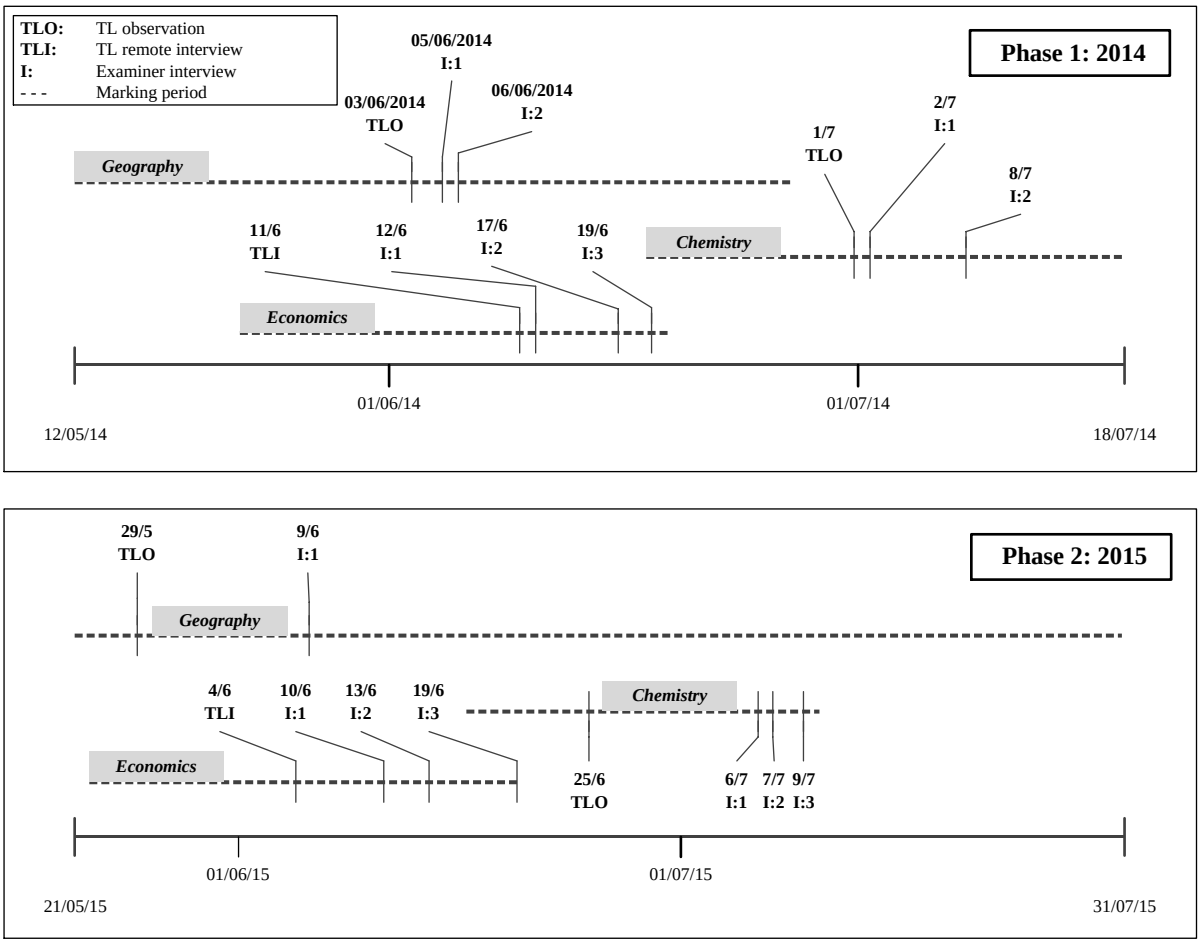

Figure 4.1: Data Collection Timeline (2014-15)

To address my first research question (What are the characteristics of examiner feedback?) I developed a data analysis approach to ensure that my study design was adequate for my intended aims. This approach set out the types of enquiry that I wanted to carry out, the levels of analysis that such an enquiry required, and the nature of the evidence collected (Table 4.3). My approach involved both deductive and inductive enquiry types, and considered discourse at different analytic levels (word, move, and message). These analyses required that I, in turn, attend to different forms of evidence in order to build a comprehensive picture of feedback discourse.

Table 4.3: Research Question 1 Data Analysis Approach

Enquiry Type	Analysis level	Evidence
1 Feedback intention (deductive)	Message and move	Content themes [Pilot codes] (Johnson, 2013)
2 Exploratory word use (deductive)	Word	Exploratory word frequency
3 Specialized word use (inductive)	Word	Keyword analysis
4 Discourse management (inductive)	Between/across message	Rephrasing, revisiting, questioning, and turn taking regularities

To structure my study, I set out to systematically capture all of the feedback messages that were given by three TLs to all of the examiners in their respective marking teams. These data included feedback that was delivered as electronic [email] messages through the *scoris*® *assessor* system or through telephone communication. This meant that the different forms of communication required that I employ an appropriate transcription approach (see below).

For my study I decided to focus on three UK Advanced level General Certificate of Education (GCE) subjects[10] (Chemistry, Economics and Geography[11]). These subjects were chosen because they included scripts that incorporated some subjective items. These items tend to invite performances that require higher order skills, and inevitably involve complex decision making on the part of an examiner when applying the mark scheme. Such items are considered to be the most complex item type, they tend to result in lower levels of examiner agreement (Massey & Raikes, 2006). Bramley (2008), and I anticipated that they would produce rich between-examiner interactions.

Discussion with OCR administrators who were responsible for employing TLs and examiners identified a group of TLs from these subjects who could be contacted. Consideration was given to whether any TLs would be intimidated or worried about being involved in a research study. The TLs who were identified by OCR were considered to be representative of usual

[10] GCE courses are usually studied over a two-year period and are widely recognised in England, Wales and Northern Ireland as being the standard entry qualification for assessing the suitability of applicants for academic courses in UK Universities. GCE students are usually around 18 years old.

[11] ***Chemistry A [Unit Code F322]***: Chains, Energy and Resources. 100 mark paper; ***Economics [Unit Code F581]***: Markets in Action. 60 mark paper; ***Geography [Unit Code F581]***: Managing Change in Human Environments. 75 mark paper.

TL practice. I gathered some background information on each TL, and this is presented in Table 4.4.

Reflecting my ambition to study feedback practices during live marking episodes, I was not responsible for sampling the examiners who were working with the TLs in each subject team. Each team was constructed by OCR, using their normal 'examiner panel' construction process. I was able to gather information about each examiner from the database that OCR administrators use to construct their examiner panels. This data was limited to gender and whether the examiner had worked as an examiner prior to the current marking session. I was able to augment this data with information from the TLs about whether they had previously worked with the examiner, as I felt that this might have implications for the common ground constructed by TLs and examiners (see Chapter 2.5). This information is contained in Table 4.4.

Table 4.4: TL and Examiner Demographic Data

Chemistry					Economics					Geography				
TL1: *Male; Age: 50s; Examiner for 20+ years; TL for 8 years; a retired teacher of Chemistry from a Comprehensive School; still continues to tutor students privately*					TL2: *Male; Age: 20s; Examiner for 5 years; TL for 2 years; currently a teacher of Economics in a Comprehensive School*					TL3: *Female; Age 50s; Examiner for 20+ years; TL for 5 years; a retired teacher of Geography from a Comprehensive Sixth Form College*				
Examiner	♀♂	Exp[1]	Fam[2]	Int[3]	Examiner	♀♂	Exp	Fam	Int	Examiner	♀♂	Exp	Fam	Int
1	F	*	*		8	M	*	*	†	14	M			
2	M	*	*	†	9	F				15	M	*	*	
3	F				10	F	*	*		16	F	*	*	
4	F	*	*		11	M			†	17	M			
5	M			†	12	M	*		†	18	M			*
6	M	*	*		13	F				28	M	*	*	
7	F				24	M			†	29	F			*
19	F			*	25	M	*	*	†	30	M			
20	M	*	*		26	M	*	*	†					
21	M			†	27	M	*	*						
22	M	*												
23	F	*	*	*										

[1] Experience (the examiner has examined in a previous marking session); [2] Familiar to the TL (worked together previously); [3] Interviewed examiner (*Face-to-face/†Virtual)

The participating TLs varied in Team Leading experience, although all had been in the role for less than 10 years. Two TLs were retired teachers (with one still practicing as a tutor) in their 50s, whilst the other was a serving teacher in his 20s.

The examiners included 11 women and 19 men. Sixteen examiners were experienced (having worked as an examiner for at least one previous marking session), and of these, 14 examiners had worked with the TL in their team in a previous marking session.

Table 4.4 also includes information about whether an examiner was interviewed as part of the project, and whether this interview was face-to-face or virtual. This is discussed further in Section 4.3.1.1 (Semi-Structured Stimulated Recall (SR) Interviews).

After gaining the informed consent of the participating TLs and examiners (Appendix F), I collated email feedback messages by remotely accessing each TL's *scoris® assessor* account. Each email message was then transferred from the *scoris® assessor* system and converted to a Microsoft Word document. At this stage I retained the formatting of the messages so that their appearance was as close as possible to how they would have looked when an examiner had encountered them. This formatting allowed me to analyse each piece of feedback at message-level. In addition, each message was tagged to include its time of delivery (date, hour, minute) to enable me to establish the chronological relationships between different messages.

In this thesis, I present the email feedback messages as close as possible to the original presentation format in the *scoris® assessor* system. Wherever possible, this includes contextual data such as the date and time of the interaction, or the subject heading of the message (although the names of participants are removed so that their anonymity can be retained). The spelling, punctuation and text manipulation choices of the participants (such as the use of bold or coloured font) are represented without any alteration. Any typographical errors found in the original messages are retained to help to signify the often time-pressured context of message construction.

I have manipulated the data in a number of ways to aid readability and analysis. The most obvious manipulation I have made is to arrange the order of messages so as to convey the chronology of the discourse. The earliest messages are located at the top of the email chain, in contrast to some email programmes that present the most recent message in a chain at the bottom of the discourse. I also indent responses to messages so that the boundary between one message and another is as clear as possible. The file name for each message is also displayed below each message. These features are demonstrated in the example below:

From: TL: 25/05/2014 16:37

Hi *Examiner name*,

7 - we kept this answer in L2. Only L1 should be awarded for a market failure diagram. We didn't give the first L3 as that is not

saying a TAX causes the private cost curve to shift. We didn't give the second L3 because the candidate is saying a tax is given to producers, which isn't right.

Hope this all makes sense. I'll send this script back to you for you to have a look at and make changes before resubmitting. Let me know if anything is unclear.

Thanks,

TL name

>From: Examiner: 25/05/2014 4:52:17 PM

hi,

Can i just ask about question 7- are you saying that entire essay was kept to level 2 as there is discussion ie pros and cons/conparison?

Examiner name

[E9 25.5 1637]

To aid the description of an analysis, I sometimes present feedback messages with line indicators. This procedure mirrors the conventions of Conversation Analysis and allows particular features of discourse to be indexed. This convention is shown in the example below:

```
TL: 18.6 1456
001   Hi Examiner name,
002   Thank you for the feedback, I have amended the 2 you sent back
003   to me.
004   2 queries:
005   ID 649581302 - Q1 g ii - MS says 'it' should be assumed to mean
006   cyclohexane. Do they still need to have written cyclohexane
007   somewhere in their answer to get the mark? I accepted 'It burns
008   more effectively'.
009   I cannot find the comment re: 'assumed to be cyclohexane' in
010   the mark scheme - it may have appeared in the practice scripts
011   by the sound of it, and was incorrect if it did.

[E3 18.6 1456]
```

In some cases I present the feedback in a condensed format so that I can highlight a particular element of interaction across the feedback messages. This approach removes any information that I consider to be extraneous to the point that I am trying to make. This sort of information could include elements of messages that discuss other items or issues that are beyond the focus of interest. This convention is exemplified in the extract below:

Examiner 29 to TL: For question 4 a(ii) if the candidate writes salt pan for a(i) and then explains how its shape has been formed using water accurately, is it possible for them to get full marks for a(ii)?

TL to Examiner 29: No because the landform is not found in this landscape. If we can we avoid double penalties but this time the PE decided not to accept these answers.

[E29 M5 25.5 15]

In the case of telephone feedback, each TL agreed to audio record telephone messages. To enable this I provided each TL with a digital recorder; all of the participants consented to this[12]. Each of the recordings was then passed to me for transcription. In contrast to the email feedback messages, where the format of the messages is largely predetermined by the technology through which they were communicated, I needed to carefully consider my approach to the transcription of telephone feedback messages. One affordance of spoken communication is that it can convey non-verbal information, with this information being important to how the communication is interpreted. For example, in addition to the actual words used, verbal messages carry additional contextual information, such as tone to convey emotion or pauses, that can give insights into the intentions and interpretations of the communicating participants. These metalinguistic elements represent the machinery of interaction (Sacks, 1963, cited in Crabtree, Tolmie, & Rouncefield, 2013) and shape its character. As a consequence, transcription needs to reflect the context surrounding the discourse as much as possible. Although data transcription is one of the earliest stages of analysis and may appear on the surface to be a relatively straightforward process, it is noted that 'transcribers must always make decisions about what to include and exclude in [their]

[12] A section on the broad ethical considerations for the research is included at the end of this chapter (4.4).

transcripts' (Bucholtz, 2000, p.1441). This is because the stance of the researcher is enacted in the transcription process (Lapadat & Lindsay, 1999, p. 66) through the selectivity of their transcription choices (Green, Franquiz, & Dixon, 1997; Swann, 2010) and through the way that they represent the participants (Ochs, 1979).

In order to capture contextual discourse features, I transcribed the TL and examiner recordings using a number of conventions that could help to display any indications of alignment or misalignment beyond the actual words communicated. These conventions were adapted from Jefferson (2004) and Martin & Rouncefield (2003) and are outlined in Appendix G. These transcription conventions allowed me to indicate features such as:

- Pause (a potential indicator of dispreferred response);
- Sigh (a potential indicator of frustration);
- High rise (a potential indicator of surprise);
- Exclamatory utterance, lengthened syllables, and increased volume (potential indicators of stressing a concept);
- Low rise (a potential indicator of disappointment);
- Overlapping speech (a potential indicator of misalignment or checking understanding).

In common with the email feedback transcription, each telephone message was tagged to include its time and date of delivery. This allowed me to analyse each piece of feedback at within-message level as well as in relation to the any other communicated messages (between-message level).

Having established the nature of the data in my study, I use the next chapter section to outline the theoretical perspective that underpins my approach to data analysis. To do this I outline my understanding of Sociocultural Discourse Analysis (SCDA) and then go on to explain how I refined this approach, integrating a number of analytical methods to create an augmented Sociocultural Discourse Analysis (ASCDA) approach.

4.2.1 Sociocultural Discourse Analysis (SCDA): Background and Methods

In Chapter 3 I set out the basis for why communication is an area of specific interest for sociocultural research. Reflecting its Vygotskian heritage, which holds that individual development is mediated through cultural tools which are developed through social interaction (Vygotsky, 1978), a sociocultural approach to the study of communication argues that it is important to recognise its interpersonal and contextualised nature. According to Lefstein, Snell, & Israeli (2015), sociocultural analyses '…seek to replace reductive units of analysis with more holistic ones' (Lefstein et al., 2015, p. 868). This, therefore, has

implications for the methods required to study it. For Gee & Green (1998), this means that 'a range of different approaches to discourse analysis [are] needed' to ensure that the interpretations generated are ethnographically grounded whilst affording a view of the complex patterns that are constructed across different interactions (Gee & Green, 1998, p. 158). Therefore, a central feature of a sociocultural analysis of discourse is a need to ensure that analysis takes account of both its particular *within-discourse* and its general *across-discourse* features.

Taking this approach further, Mercer (2004) and Littleton & Mercer (2013) have discussed the need to use sociocultural discourse analysis (SCDA) as a framework to explore talk in a variety of learning situations. According to Mercer, SCDA can be distinguished from other approaches 'because it is based on a sociocultural perspective on the nature and functions of language, thinking and social interaction' (Mercer, 2004, p. 138). This claim reflects two key sociocultural themes: firstly, that knowledge is created through social interaction and communication; and secondly, that interlocutors reflexively create their own context of common knowledge through the process of interaction.

The SCDA approach developed by Mercer (2004) and Littleton & Mercer (2013) provided the broad frame for my analytical approach. In line with the observations by Gee & Green (1998), SCDA seeks to investigate the particular and general features of interaction. In real terms this means that the analytical structure for enquiry is aimed at the levels of exchange, utterance and word. This multi-level approach allows the social intentions and purposes of interaction to be considered alongside its structural organisation. For me, SCDA can be distinguished from other forms of discourse analysis by its particular interest in four themes. SCDA focuses on lexical content and the cohesive structures of talk that are utilised by participants interacting in a specific cultural context (e.g. a workplace). It is also specifically interested in how shared understanding is developed in social context and over time (Mercer, 2008b), as well as being concerned with the impact of talk. SCDA also accentuates that way that discourse is used to pursue *Joint Intellectual Action* or interthinking (Littleton & Mercer, 2013). Based on this work, and as stated earlier, my analytical approach to SCDA seeks to evidence four themes:

- Content - the lexical content and the cohesive structures of discourse;
- Time - how shared understanding is developed in social context over time;
- Joint Intellectual Action - evidence of how participants recognise the thinking of other participants and use this to coordinate shared understandings;

- Impact - the effect that discourse has on the cognition and behaviour of the participants, which in the case of examiner feedback is to attain a point of agreement or resolution.

For the purposes of my study I refined the original concept of SCDA, highlighting some specific areas of interest for my study. Building on the emergent picture of feedback interaction that I developed in my pilot study, I propose the use of a mixed methodology that incorporates three approaches to explore the four sociocultural themes that pertain to the original SCDA approach. As a result, I call my methodological refinement 'augmented SCDA' (ASCDA), and it comprises Thematic Content Analysis, which is a strand of Discourse Analysis (DA), Conversation Analysis (CA), and Corpus Linguistics (CL).

4.2.2 The Discourse (Thematic Content) Analysis Approach

From a sociocultural and realist perspective, thematic analysis was a useful approach for analyzing my data. Thematic content analysis (TCA) is a 'respected… well-established' (Boje, 2001, p. 122) and 'widely-used' (Braun & Clarke, 2006, p. 101) approach in Discourse Analysis. Content analysis is a general term that covers 'a number of different strategies for analysing text' (Vaismoradi et al., 2013, p. 400). For example, according to Holsti (1969), 'content analysis is any technique for making inferences by objectively and systematically identifying specified characteristics of messages' (Holsti, 1969, cited in Franzosi, 2004, p. 187). Braun & Clarke (2006) argue that content analysis is assumed to be a relatively unproblematic method to use from a realist perspective because language represents units of conveyed meaning which reflects, and enables us to articulate, meaning and experience. In addition, Braun & Clarke (2016) go on to highlight that 'pre-existing' themes are not 'identified' by researchers, but that they are constructed by researchers based on their interpretative choices (Braun & Clarke, 2016, p. 740).

As I outlined at the beginning of this chapter, I also acknowledge that a sociocultural perspective highlights the role of social context on interpreting discourse, and that this can represent a challenge for researching communication. Similar to some constructionist arguments (e.g. Burr, 2003; Pinkett, 2000), a sociocultural perspective argues that the meanings constructed by discourse participants are embedded in the social and cultural activities that they mutually engage in (e.g. professional discussion in the workplace), rather than inhering within individuals. This perspective is supported by Sawchuk & Stetsenko (2008), who note that social facts are created via practices mediated by cultural and material resources which interlink local-level enactment with extra-local control. This means that the

process through which I interpret such communication is also contextualized, and that theme identification is interconnected with my own perspective. The role of researcher subjectivity in thematic analysis is a concern for some (Guest, MacQueen, & Namey, 2012, p. 15). This means that I needed to take steps to assure the quality of my analyses. One of these steps involved me being explicit about my perspective and the theoretical framework on which my thinking was based. This explicit referencing helps the plausibility of the outcomes generated by my analyses to be evaluated against a separate body of knowledge and its theory base. Such a concern also led me to consider the approaches that researchers can adopt to quality assure their research outcomes (see reference to Braun & Clarke, 2006, below for further discussion).

A central element of my study design was to collect and analyze a large corpus of feedback texts from six examining teams (across a pilot phase and an additional two year period). I argue that this scale of data collection supported my interpretative process because it allowed me a privileged view of the discourse across a number of cases. This perspective had not been possible prior to this study. My experience of working in a professional assessment environment also allowed me to draw on knowledge (e.g. the meanings surrounding specialised linguistic terminology) that cohered to some extent with those participating in the research. This is important as it is claimed that a goal of qualitative research 'is to arrive at an understanding of a particular phenomenon from the perspective of those experiencing it' (Vaismoradi et al., 2013, p. 398).

I used a version of content analysis called Thematic Content Analysis (TCA) for the pilot phase of the project, and then used this as a foundation for the subsequent project phases. TCA is a version of Discourse Analysis (DA) and it allowed me to consider the feedback texts in a top-down way. Such an approach allowed me to take into account the context around the text, as this was expected to influence the meanings constructed by the discourse participants. The purpose of TCA is to identify themes or patterns across a data set which capture and exemplify important elements of the data. According to Gibson & Brown (2009), this analysis allows commonalities, differences and relationships across a text to be explored. When carrying out TCA it is important to define what is meant by a 'theme'. I employed the general description used by Guest, MacQueen, & Namey (2012), who state that 'a theme is a unit of meaning that is observed (noticed) in the data by a reader of the text' (Guest et al., 2012, p. 50).

In the following section I outline how I developed my codes and overarching themes for my TCA. I then outline the codes and how they relate to my theoretical perspective.

The TCA process relies in the first instance on coding. Coding involves the systematic application of codes to a text in order to extract uniform and standardized data. These codes capture the dominant ideas, or constructs, in a text through 'encoding qualitative information' (Boyatzis, 1998), and are usually words or phrases that serve as labels for sections of data. These codes are the conceptual foundations of a theme, with coding preceding the process of theme generation. Themes represent a less abstract phase of analysis than the preceding coding phase (Guest et al., 2012), since they bring together fragmented and distributed codes into a coherent grouping. The coding framework that I generated in my pilot study provided a foundation upon which to build a refined coding framework for use with the larger data set that was gathered in the subsequent project phases.

My TCA approach employed a largely inductive procedure, as I was exploring the text for themes that could explain how examiners aligned their thinking. In this way I hoped to ensure that any themes that I identified were 'strongly linked to the data themselves' (Braun & Clarke, 2006, p. 83). At the same time, I recognised that the theoretical framework within which I was working, and that helped to focus my research, predisposed me to noting particular features. Braun & Clarke (2006, p. 87) outline phases of analysis that should be carried out by a researcher in order to assure the quality of their analysis: these phases include familiarisation; code generation; codebook refinement; theme generation; and theme definition. I outline each of these phases below, along with a commentary on how I applied them in my analysis.

Familiarisation: Firmin (2008) suggests that multiple re-reading of the text helps with this process. In my study, this familiarisation phase started with the manual data transfer from the examiners' digital marking system to my own data repository. During this time I read each message briefly as I moved it to an appropriate file space. I also needed to transcribe any telephone discourse as part of this data handling process. Once each message had been isolated and filed, I inputted it to the analysis software. This process required me to read each message carefully and to start to notice elements of similarity and difference in the content covered in the sentences or phrases of messages. This led to the second TCA phase.

Code generation: This is a comparative process that relies on the researcher relating the ideas presented throughout the text. This results in the first draft of a coding scheme, where code

meaning is defined in terms of code categories, types and relationships. Boyatzis (1998) notes that the development of a codebook helps support the consistent application of codes through labelling, defining, and describing codes. This is important because consistency can be an indicator of the robustness of the classification system (Fereday & Muir-Cochrane, 2008).

I used a memo function in the computer-aided qualitative data analysis software (CAQDAS) system to generate the codebook. Reviewing the use of software in qualitative analysis, Lu & Shulman (2008) note that the use of CAQDAS can improve analysis by helping to make the examination of data more complete and rigorous, although Kvale (1996) cautions against any tendency to lose sight of the context in which the data is situated through the use of such technology. My coding/memo system recorded the name of each code along with a working definition and exemplars. This allowed me to revisit my assumptions as I moved through the data coding phase. My coding scheme included some low-inference codes (e.g. noting opening and closing indicators or noting the instances where information was being given), as well as some higher inference codes. These higher inference codes included indications of the types of information being given, and the types of linking and referencing moves that were evident in the text.

As a quality assurance (QA) check on the consistency of my coding application I carried out a recoding exercise. This involved double coding a sample of 54 randomly selected messages. This recoding amounted to 9.7% of the total sample and conformed to the amount used in other studies that have analysed qualitative data (e.g. Renz, Watts, & Conrad, 2013; Salerno et al., 2002; Suárez-Orozco & Hernández, 2012). Dual coding allows for the identification of particular codes where there is a lack of clarity.

To establish the extent to which my coding application was consistent I used a number of measures. At the most basic level I calculated the percentage agreement of code application across the sample. Percent agreement (also called simple agreement) is a popular measure of agreement but, according to Joyce (2013) there are a number of weaknesses related to this measure. These weaknesses include there being no comparative reference point to relate the rate of agreement to indications of chance, and the measure hiding important disagreements as it averages across multiple cases, so the outcomes are potentially biased by including variables with low variance and high agreement levels. Despite this, it is widely used and ‘Coefficients of .90 or greater are nearly always acceptable, .80 or greater is acceptable in most situations, and .70 may be appropriate in some exploratory studies for some indices’

(Neuendorf, 2002, p. 145). Joyce also notes that for social science studies in the communication field, the goal is often .80 pairwise agreement (Joyce, 2013).

To augment my consistency check I carried out some additional tests. For more sophisticated analysis of agreement across a variable with two values, Joyce (2013) and Lombard, Snyder-Duch, & Campanella Bracken (2010) suggest that Scott's π, Cohen's κ, and Krippendorff's α are appropriate measures. This is because they improve upon percent agreement by factoring in the extent to which a given value will be coded by chance (Scott's π, Cohen's κ), or by measuring observed and expected disagreement (Krippendorff's α). Because of this, these measures are considered to be more conservative than measures of percentage agreement, and so interpretation of their outcomes would suggest that coefficients of .80 or greater is acceptable in most situations (Lombard et al., 2010). This is reinforced by Krippendorff (1980) who notes that for α 'it is customary to require $\alpha \geq .800$' (Krippendorff, 1980, p. 241).

The recoding analysis showed that my coding categorisation was good (Hruschka et al., 2004) (Table 4.5), but that the coding of *Locating Credit* and *Rationalising Credit*, and *Marking Principles* were less reliable than others, and required the memo system to be refined to describe the codes more clearly prior to additional coding.

Table 4.5: Sample Double Coding Consistency Outcomes

	Phase 1 Re-code *(10% sample/n=54)*	**Phase 1/Phase 2 Re-code** *(5% sample/n=28)*
N Agreements	26	27
N Disagreements	5	5
N Cases	31	32
N Decisions	62	64
Percent Agreement	83.9	84.4
Scott's π	0.83	0.83
Cohen's κ	0.83	0.83
Krippendorff's α	0.83	0.83

As my analysis was being carried out over two phases it was important for me to carry out a re-standardisation exercise prior to the coding of the second phase of feedback data. Standardisation allowed me to reflect on my current coding judgements in respect to my earlier coding decisions. This activity was important for two reasons. Firstly, a measure of reliability has a QA function, helping to establish that the explicit rules of coding are being systematically applied to the whole dataset and, therefore, that my understanding of the coding scheme was consistent over the two phases of the project. In this way I could demonstrate that I had a reasonable amount of judgement-making objectivity during the

coding process (Frankfort-Nachmias & Nachmias, 1996, p. 324). The second purpose of re-standardisation was conceptual. According to the work of Laming (2004), the application of human judgement requires a return to a standard reference. This is because judgements are considered to have a relative basis and are prone to shifts over time if they are not related to previous decisions. By returning to my previous coding I was able to recalibrate my judgement-making ability prior to engaging in the Phase 2 coding exercise.

My re-standardisation exercise involved recoding a 5% sample of the original coded data (i.e. 28 messages), and to undertake more standardisation using a 10% sample if there was poor reliability (i.e. a kappa coefficient which was in the range 0.60-0.74 (Cicchetti, 1994)). The initial recoding outcomes showed that my coding reliability was good. A result of this exercise was to collapse the *Orientation* code into the *Bridging* code in the original dataset. I also further refined the definition of the *Balancing Perspectives* and *Request Information: Monologue* codes to better differentiate them from other codes.

Codebook refinement and theme generation: The refinement phase involves the application of the codes to the data set, which continues until a point where saturation is reached and new codes are no longer generated. In keeping with the pilot coding convention, I employed a double coding approach (for the reasons outlined in Chapter 4.1.2). In my analysis, this saturation point occurred roughly half way through the analysis of the second TL's case. I was also able to amend and sometimes collapse codes into subsidiary themes where they seemed to overlap but demonstrated some nuanced differences. In this analytical phase, the codebook becomes a resource for collating codes into themes, as it allows the researcher a view across the texts using a common set of descriptors. A code mapping exercise can aid the generation of themes as it is possible to see how codes interact with each other and allows the associations between codes to be conceptualised. In my study, themes were associated with the functions that the participants were trying to achieve through their discourse. I was able to collate the codes into several categories. These categories included *Non-Feedback Information*; *Interpersonal Information; Giving Feedback Information;* and *Actions*.

Conscious of the concerns about the interpretative role I played in the coding process, I also adopted an additional QA stage to those reported by Braun & Clarke (2006). To augment the QA process, I carried out a member check stage by passing my coding and examples to the TLs. The reflections of the TLs helped to reassure me that my code interpretation was valid. In addition, I was also able to review the themes by checking on potential overlaps and

interactions. In my analysis I was able to use the temporal organization of communication to shape the way that the themes were outlined (e.g. themes covering the stylistic elements of messages preceded themes covering the information conveyed, which then preceded themes that covered the outcomes that resulted from communication).

This organization framework also allowed me to check whether there were areas of overlap between the themes. Again at this stage, re-reading of the codes within the themes allowed me to consider whether the codes accurately represented the content contained within them. This led to the final stage of analysis which was *theme definition.* This stage involves data reporting where vivid examples of selected extracts illuminate the themes being described. At this final stage the researcher looks for congruence between the extracts chosen and the analytic claims being made.

My analysis was not a purely inductive coding process. I had a deductive steer, drawing on a loose theoretical framework that was generated during my pilot study. This meant that I was, to an extent, already looking for evidence of hypothetical features that I had observed in my pilot study. My deductive framework focused on the content and the qualities of the communication conveyed. My thematic analysis led to the refinement of the coding framework that I generated during the pilot study. The coding process also allowed me to tag the codes to the item and mark scheme type features of the feedback messages. This enabled me to consider any associations between codes and item/mark scheme features that had been hinted at in my pilot study.

There was also space in my analysis for inductive analysis. According to Thomas (2006), inductive approaches are primarily based on the use of detailed readings of raw data to derive concepts, themes, or a model through interpretations made from the raw data (Thomas, 2006, p. 238). In my analysis there was an opportunity for the raw data to inform the generation of new codes. This was supported by my adoption of a CA inspired approach within my analytical process (see below). This approach involved continually interrogating the data for themes by asking ‘why is this happening now, and to what purpose?’ This iterative reflection is a guiding principle for CA approaches and involved initial close reading, and revisiting of, the text as the coding process was refined. I was also able to use the MAXqda CAQDAS system to consider how themes interacted with each other which, according to Kvale (1996), is sometimes referred to as theory building.

4.2.2.1 My Coding Framework

My final coding framework included 4 superordinate and 30 subordinate codes (Appendix H). 13231 codes were assigned to my transcribed texts, comprising of 8058 tag codes (i.e. item/mark scheme features) and 5173 discourse codes. 2138 items were coded in the feedback communication.

My coding structure was organised into four themes, broadly reflecting the SCDA themes of interest that informed my study (these are set out in Table 4.6, and outlined more fully in Appendix H). Two of the superordinate themes dealt with *Content*, looking at this from its transactional and its interactional functioning (Brown & Yule, 1982), one theme considered *Joint Intellectual Action*, and another theme considered *Impact*. The SCDA theme of *Time* could only be considered in terms of each of the other themes.

The set of codes that were clustered into the superordinate code/theme of *Giving Information* were most clearly related to the types of *Content* being communicated between the TL and the examiner, and how language was being used to perform a transactional linguistic function (Brown & Yule, 1983). These codes related specifically to marking concepts (e.g. the location or reason for marks), and to technical information (e.g. about how to use specific mark-related tools such as annotations). Some codes in this themed grouping were also able to capture something of the intentions underpinning communication through considering whether the communication was invited (e.g. a request) or enforcing (e.g. a directive).

The set of codes that were clustered into the superordinate *Focus/Interpersonal* theme related to stylistic discourse elements (Holmes, 2001). These elements included *Content* that was non-feedback specific and largely interactional in function (Brown & Yule, 1983). These codes indicated those discourse elements that were, to an extent, content-free and had no standalone substantive meaning attached to them in terms of conveying new information to a recipient. This also meant that these codes often required double-coding (e.g. with an additional *Giving Information* code) to aid interpretation. These codes included openings and closings, indications of agreement and disagreement, politeness and authority markers, and accentuation. I also included references to social common ground in this section as this element was not necessarily linked to feedback content per se.

Table 4.6: My TCA Themes and Codes

SCDA Theme	**TCA Superordinate Theme/Code**	**TCA Subordinate Code**	**Explanation**
Content	Focus/ Interpersonal [FI]	Opening	*Start of the interaction*
		Closing	*Closing of the interaction*
		Disagreement	*Conveying disagreement*
		Agreement	*Conveying agreement*
		Confronting gaps	*Dealing with different understandings*
		Distancing	*Positive and negative politeness*
		Authority	*Allusion to power/senior examiner team*
		Social common ground	*Reference to background knowledge of each other*
		Accentuation	*Highlighting (e.g. exclamation marks)*
Content	Giving information [GI]	Directive information	*Outlining what to do next*
		Marking principle information	*Information on a general marking principle*
		Privileged information	*Senior Examiner only information*
		General information	*Non-specific information on particular marking*
		Technical information	*Marking system level information*
		Mark statement	*Number only/statement of mark*
		Location of credit	*Limited perspective: pointing out that there is a difference*
		Rationale for credit	*Extended perspective*
		Standards	*Insight into standards*
		Examiner rationale	*Giving examiner point of view*
		Request information (Closed)	*Request for definitive information*
		Request information (Open)	*Request for non-definitive information*
Joint Intellectual Action	Bridging [Br]	Internal reference	*Reference to content within same message*
		External reference	*Reference to content across different documents*
		Historic reference	*Reference to past and present content*
		Cultural reference	*Reference to content in another context*
		Balancing perspectives	*Recognising ambiguity*
		Offer support	*Making support role evident*
Impact	Action [Act]	Review	*Indication of review needed*
		Resolution	*Indication of agreement with other's perspective*
		Reification	*Examiner adding notes to mark scheme*

The set of codes that were clustered into the superordinate *Bridging* code/theme covered discourse elements where links were being made across different sources of information. Some of the codes in this section signified where exophoric and endophoric relations (Brown

& Yule, 1983) were being used by the participants to link important information across time (e.g. with reference to past shared discourse) and across texts (e.g. different text sources such as commonly available mark scheme documents). This section also included codes where a participant directly referenced the perspective of the other discourse participant (i.e. a balanced perspective being a signifier that one person could recognise the other's thinking), or where support was offered (suggesting that an identified need was being recognised). All of these codes implied that the participants were considering each other's perceived needs when crafting the discourse. This indicated a form of *Joint Intellectual Action* (as represented in the SCDA framework) as one person needed to construct and hold an idea of the other in their mind as they prepared their discourse. The codes in this section were also most likely to reference the *Time* dimension of the SCDA framework as there was reference to elements of past interactions.

The set of codes that were clustered into the final superordinate *Action* code/theme covered discourse elements where there was reference to something that was done as a result of the feedback. These codes most clearly linked to the *Impact* dimension of the SCDA framework and included references to participants' reviews of their own work as a result of feedback, indications of agreement (that served as a signifier of the closure of a disagreement), and the annotation of new information by a participant as reified information (as this signified the creation and recording of new common knowledge).

The scale of the thematic coding allowed me to carry out some analyses using descriptive statistics. These analyses allowed me to identify which relations were most common (e.g. where there were links between historical references and resolving disagreement), and then to explore more unusual cases (e.g. where there was reference to any past working relations between participants). The close reading process underpinning the thematic analysis process also enabled me to identify keywords for exploration via a CL approach. For example, the term 'again' appeared to be an indicator of where examiners linked ideas in their text to earlier references within their shared discourse.

My use of methods associated with TCA contributed to my ASCDA methodology by enabling me to construct an overall shape of the large corpus of feedback data, whilst also allowing me to analyse the way that the feedback data (as represented by my coding) changed over time, which was an important element of my theoretical framework. Rigorous coding, drawing on a mix of inductive and deductive interpretation processes, gave insight into who

communicated what, when, and how. In the next section I will outline how methods associated with CA contributed to my methodology.

4.2.3 My Adapted Conversation Analysis Approach

Although I sought to gain new insights from the use of a broadly CA perspective, my methodology involved some adaptation of conventional CA methods. According to Ayaß (2015), transcribing data is a form of re-work which makes it visible for analysis. This process is not theory neutral because the choice of approach influences the data to be analysed. Bearing this in mind, I chose to use an adapted version of CA transcription conventions so that I could focus on elements of how shared understanding and repair were evident in the participants' spoken utterances. This is outlined in more detail below.

I integrated a non-standard CA approach into my analysis in a number of ways:

- I chose not to build my analysis empirically from the spoken data alone, instead using data from written/text-based sources;
- I drew on theoretical literature and data from other studies to explore my spoken data for the prevalence of 'Exploratory Talk' in a deductive fashion;
- I used CA to gather evidence that reinforced analyses based on other, more quantitative methods. Stivers (2015) observes that the notion of quantitative coding is considered by some to be antithetical to the spirit of CA, although it is noteworthy that others suggest that this is not problematic if the nature of the phenomenon of study requires it (Hoey & Kendrick, 2017).

Two themes that became salient from the data familiarisation phase were explored further through CA tools. One theme related to the use of politeness strategies (e.g. the use of apology), which suggested that language was used to support a cumulative function, and to build and maintain on-going engagement between participants. The second theme of note was that participants were using language to locate themselves socially in relation to their participants. This point shares some similarities with Positioning Analysis (Bamberg, 1997), which explores how discourse construction reflects a speaker's perspective, including how they position themselves with regards to their audience. In this way, the outcomes elicited through the use of CA-inspired methods were informed by the nascent themes that emerged from data familiarisation, and were therefore to some extent already theory laden.

The second level of my analysis sought to steer away from a theory driven approach by using CA tools to identify patterns in those data where there was direct interaction between participants (i.e. telephone feedback or email messages where there was embedded text

containing both participants' input). According to CA, recipient design (Sacks, 1992) is a fundamental principle for the production of talk. This principle suggests that speakers orient themselves to recipients when they construct discourse. I also extended this notion to my analysis of two-way text-based feedback discourse. My rationale for this was that email discourse is often, like talk, structured around paired-unit communication patterns. This position is supported by other studies which have shown that email communication can possess informal characteristics that are similar to those found in verbal communication (Georgakopoulou, 2011; Waldvogel, 2007). Another influence on analysis was the observation that one of the most fundamental and well attested patterns in conversation is a 'preference for agreement in the second turn of an adjacency pair' (Myers, 2004, p. 113), and that disagreement in conversation tends to be delayed, modified, or prefaced (Pomerantz, 1975, 1985; Pomerantz & Heritage, 2012; Sacks, 1987). Since a defining feature of the feedback genre is considered to be that it involves managing bad news (Yelland, 2011; King et al. 2008), it would be anticipated that navigation around negativity, whilst ensuring articulated activity, would also be a feature of feedback communication. I interpret the notion of articulation work (Strauss, 1985; Schmidt, 1994, 2011) as describing the mechanisms through which TLs ensure that work across a group of individuals interconnects and reaches its intended aims. This interpretation is also broad enough to include the interpersonal skills that allow TLs to 'develop bonds' (c.f. Carpenter & Wisecarver, 2004) with relevant examiners in order to ensure job completion.

ASCDA differs from CA in its approach to transcription. Schegloff (1997) argues that a detailed analysis of talk must take place to ensure that a researcher's ideological, political or theoretical gloss is not superimposed onto the interaction event. Precise transcripts of interactions help to ensure that the participants' orientations are demonstrable, although there are criticisms that researcher selectivity still exists around the choices of interaction episodes used to support analytical claims (Wetherell, 1998). I outlined earlier how my approach to analysis had a deductive steer, predisposing me to a certain extent to look for particular discourse features. Therefore, my approach to transcription involved using a more limited annotation pallet (compared with CA transcription protocols) as I sought to evidence some particular features of importance to my study context (e.g. markers of dispreference and disagreement).

The ambition to gather and analyse naturalistic phenomena is a central principle for both CA and ASCDA, and this was also an important element of my study. As a result, I captured

evidence of naturally occurring feedback interaction by creating a corpus of examiners' online and telephone messages. Email data was perhaps least problematic in terms of gathering and processing for purposes of analysis. These data were pre-structured, with the initiator of the communication having a great deal of control over its physical appearance before passing it to the recipient. The asynchronous nature of email messages also simplified some of the usual complexities around transcript generation. Email messages have linear characteristics, with adjacency pairs (Schegloff & Sacks, 1973) being physically organized in extended turn units. This linear organization imposed a natural chronology of developing discourse and can be usefully reflected in a transcript that employs a 'top-bottom' bias. It was therefore appropriate to organize the email transcript for analysis in a way that was faithful to the physical layout of the original text as a record of messages produced and received.

For the reasons that I outlined earlier, I needed to carefully consider the transcription of examiners' telephone feedback communication. Evidence from the pilot study showed that telephone feedback conversations involved initiation-and-response patterns. These interactions were generally asymmetrical, being steered by the TL's perspective (Johnson, 2013). Top to bottom sequential transcription helps to illustrate the pattern of turn units within interactions, and allows links to be made across and within texts.

Spoken feedback included intonation and prosodic qualities that could give insights into how understanding was being built and maintained. Ochs (1979) and Martin & Rouncefield (2003) suggest conventions that might be useful to illustrate such qualities; these include: '?' marking a high rise; '!' marking an exclamatory utterance; 'CAPS' marking increased volume; '_____' marking stress; '[ ]' marking overlapping speech, and ':' marking lengthened syllables. These complement the conventions set out by Jefferson (2004), and support the analysis of elements of interaction structure that are central to CA approaches.

Since transcription forms the earliest stage of analysis, it is also possible there is a need to provide mark-up tools that show meta-transcription involving my own overlaid analytical comments. The example below (Figure 4.2) shows how I overlaid my analysis onto a transcript whilst trying to minimise interference with the discourse flow:

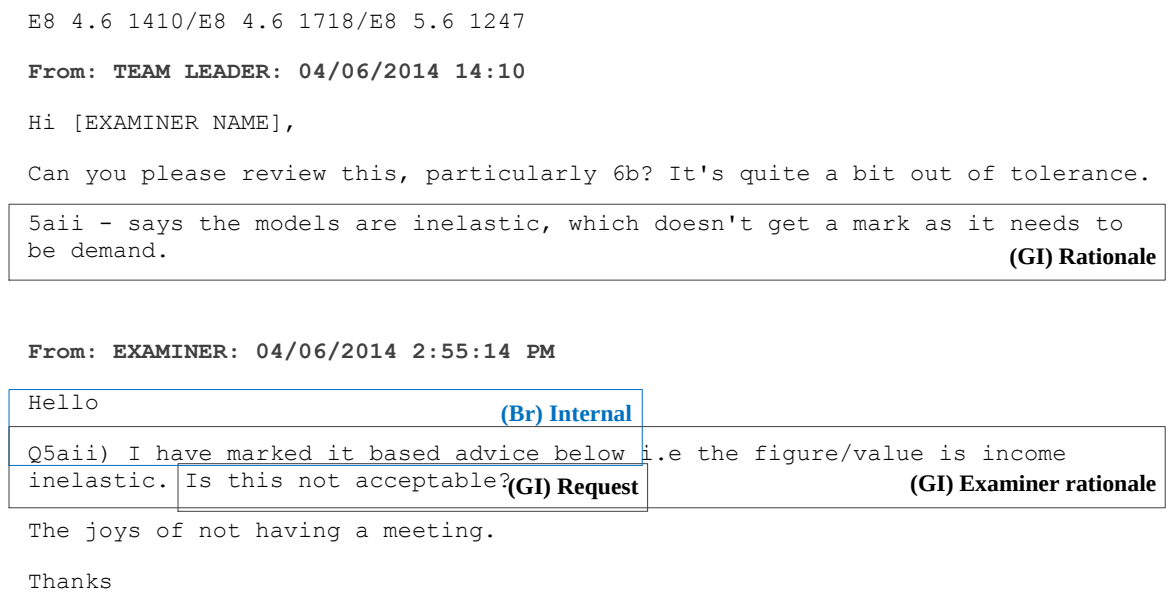

Figure 4.2: Overlay of Codes onto Feedback Text

Bearing in mind the trade-off between transcript access and legibility, I used the transcript conventions sparingly, with decisions about their inclusion being guided by the extent to which they illustrated misalignment or meaning clarification.

The use of CA tools allowed me to explore the transcripts through a process of iterative reflection, using the principle 'why that, in that way, right now?' (Schegloff & Sacks, 1973, p. 299; Seedhouse, 2004, p. 16). This analysis showed that participants used some common strategies to mediate potentially strong disagreement and loss of face. These strategies included the use of Upgrading and Downgrading (Pomerantz, 1985) and Positive and Negative Politeness[13] (Brown & Levinson, 1987). As I have discussed, politeness has links with threats to participants' self-esteem and 'face' and has been observed in other feedback contexts (Sussman & Sproull, 1999). Goffman's (1955) concept of face work suggests that identity is projected in the way that professionals present themselves. Face is the social value a person claims in an interpersonal contact (Goffman, 1955), and actions which invalidate their identity (e.g. performance failures) lead to embarrassment. Goffman claims that this is a

[13] Positive politeness reduces the threat to the recipient's positive face by accentuating empathy and common ground between the participants. Negative politeness avoids imposition on the recipient's negative face (i.e. the desire to act unimpeded) by creating respectful distance.

problematic situation since embarrassment interrupts social interactions, which face work, being given or received, has a role in diffusing.

A criticism sometimes levelled at CA is that it generates informal and imprecise analytic methodology categories 'so that they are difficult for others to use in any practical way' (Brown & Yule, 1983, p. 231). In order to counter these criticisms, Heritage (2004) has formulated a framework for interpreting the way that communication functions in social institutions (Table 4.7). I used this basic framework to provide a consistent basis for analysis across my verbal data. One useful feature of this framework is that it focuses on interaction turn elements. This approach allowed me to consider the reactions to utterances and to identify preference or dispreference, as these can be indicators of the degree of alignment attained by the participants.

For CA, the specific information communicated in an interaction is less important than the pattern of utterances. This generally runs counter to the focus of ASCDA which explores how content is used and re-used as it comes to constitute common ground for participants. At the same time, I would argue that Level 5 of Heritage's Analytic Framework (Lexical Choice) allows space for the consideration of how content functions within communication.

Table 4.7: An Analytic Framework for Structuring CA (Heritage, 2004)

Level	Name	Descriptor
1	Turn-taking organization	Question-Answer organization
2	Structural organization	Phases/sections, topics of interaction
3	Sequence organization	Opening up, developing, and closing actions
4	Turn design	What does the discourse achieve? What is the effect?
5	Lexical choice	Words of interest; euphemisms
6	Forms of asymmetry	Participation balance, Access to knowledge

My use of methods associated with CA contributed to my ASCDA methodology by enabling me to gather insights into relational interaction around turn-exchange points (e.g. through expressions of preference and dispreference). This analysis highlighted the ways that the participants used strategies, such as politeness, to mediate disagreement and to preserve professional face. In the next section I will outline how methods associated with CL contributed to my methodology.

4.2.4 My Adapted Corpus Linguistics Approach

In conjunction with the use of CA tools, I also chose to use elements of a CL approach to support my data collection. I chose not to build my analysis empirically from the use of CL

methods alone; instead I used concordance methods to augment my analysis that were also based on TCA and CA approaches. I also drew on theoretical literature and data from other studies to explore my data for the prevalence of 'Exploratory Talk' in a deductive fashion. I used a CL-based approach to analyse my feedback data at two levels: a keyword search to identify words that were used significantly more by TLs than examiners, and a targeted search for specific reference and exploratory words that were informed by the other ASCDA methods.

My analyses drew on concordance and collocation analysis methods that are commonly used in CL study. Concordance analysis can be used to gather all of the instances of potential signifiers of exploratory discourse and to present them in context as a string of co-occurring words. Concordance is a precursor to collocation analyses. Collocation is 'a psychological association between words […] up to four words apart and is evidenced by their occurrence together in corpora more often than is explicable in terms of random distribution. This definition is intended to pick up on the fact that collocation is a psycholinguistic phenomenon' (Hoey, 2005, p. 19). Collocation can give insights into the way a particular word is functioning within a text, and is possible because of the recognition in linguistics that there is a strong connection between word meaning and the syntactic structures in which word meaning occurs. Due to this relationship, the patterning of language is an important cue for how to interpret word use. In the next section I outline how I carried out my keyword and targeted search analyses.

4.2.4.1 Keyword Search

My first level of analysis involved a comparative analysis of the feedback corpus with other corpora. This comparative process highlights features that are particular to a discourse type and results in the identification of keywords. For this purpose, I identified two corpora to use as comparators with the feedback corpus. These corpora were taken from the BNC-Baby v2 corpus (BNC Consortium, 2003), which is a sampling of the British National Corpus (BNC Consortium, 2007)[14]. BNC-Baby v2 contains four million words and represents four different sub-corpora. For the purpose of my study I chose two of the BNC-Baby sub-corpora. The Academic Subcorpus contained 30 titles randomly selected within different subject areas and identified as 'written academic'. These included data originally published in periodicals and

[14] The British National Corpus (BNC Consortium, 2007) is a large corpus that is a snapshot of British English at the end of the twentieth century.

books and represents language that is being used to communicate subject content. This transmissive function shares similarities with a key characteristic of feedback discourse, which my analyses suggest involves the communication of judgements around content (Johnson, 2013). The second sub-corpus contained unscripted conversation data. This conversational data consisted of 30 randomly selected texts of conversations involving participants considered to be broadly representative of the UK population in terms of age, gender, region, and class. This conversational corpus was relevant to my study because TLs and examiners at times choose to communicate feedback via spoken telephone discourse, a practice first observed in my earlier study of examiner feedback (Johnson & Black, 2012b).

I subdivided the feedback corpus into various sub-corpora to aid comparative analysis within the feedback corpus itself, and this allowed me to generate TL keywords. These are words that are more frequently used by TLs compared with examiners in the feedback corpus. These sub-corpora are shown in Appendix I. Prior to analysis, each of the different corpus texts was converted to .txt files and loaded into Monoconc Pro 2.2 software (Barlow, 2012). In addition to this, the Stop List (words that were not to be included in the text search) was carried forward from the pilot phase and loaded into the software. This Stop List is shown in Appendix E.

This search identified words that were relatively overused by TLs (i.e. words that were significantly under-represented in any corpora except the TL corpus). According to Baker, 'a word is key if it occurs in a text at least as many times as a user has specified as a minimum frequency, and its frequency in the text when compared with its frequency in a reference corpus is such that its statistical probability as computed by an appropriate procedure (e.g. …log-likelihood score, or the chi-squared test) is smaller or equal to a p value specified by a user' (Baker, 2004, pp. 346–347). I defined TL keywords as those words that were found in a significantly greater number (i.e. signified at the 5% confidence level) in the TL sub-corpus than in each of the examiner, BNC academic, and BNC spoken sub-corpora. I also set the specification that these words needed to account for more than one in a thousand words of the total corpus (i.e. 89 words), to ensure that there was a degree of commonality about their use across the discourse texts.

4.2.4.2 Targeted Search

My second level of analysis varied to some degree from traditional CL studies. Here I used CL methods to carry out a targeted search to explore how particular words were used in the

corpus. For this process, I selected words that I had reason to consider as being of interest to my enquiry based on either insights gleaned from my ASCDA work, or on previous theory. This targeted search involved initially identifying words that might have a particular function in examiner feedback, and using CL tools to explore the characteristics of their use across the sub-corpora.

Search terms were located through two sources. Through other elements of my ASCDA approach I had noticed how specific referencing terms were being used to potentially relate different pieces of content information, or to implicate the TL and examiner relationship. These terms appeared to convey important contextual information because they implied relations between participants (e.g. in the way that a TL acknowledged the perspective of an examiner in a phrase such as 'I can see why *you* gave this mark'). These referencing terms were also being used by participants to establish exophoric and endophoric relations (Brown & Yule, 1983, p. 192). These terms appeared to help the participants to link important information so that they could convey their understanding to each other. These terms are shown in Table 4.8.

Table 4.8: Contextual Reference Terms from the TCA and CA Phases

Word	**Description of potential function**
Again	Exophoric reference: using language to make historic links
Here/They	Endophoric reference: using language to link elements within a text and ensure common focus
Me/My/We/You/Your	Personal reference: using language to bring together perspectives

If the use of contextual reference terms represented an empirical grounding for the search, the second source represented a theoretical grounding for the search. My second source was the list of potential signifiers of 'Exploratory Talk' that were taken from Wegerif (1997), and were used in the pilot study phase (Table 4.9).

Table 4.9: Potential 'Exploratory' Terms from the Pilot Study Phase

Word	**Description of potential function**
Because/Thought/Why	Explanation
But	Clarification and emphasis
For/Would	Identifying creditworthy features
If/Should/Then	Rule transmission; Conjecture
So	Verification; Rule transmission; Explanation
Think	Alleviate forcefulness; Explanation

In the next section I describe how I address my second research question, 'What are the characteristics of effective examiner feedback?'

4.3 Main Study Research Question 2: What is Effective Feedback?

To address Research Question 2, I identify dimensions that influence feedback effectiveness. This process involved case study interviews with a sample of TLs and examiners. Table 4.10 outlines my data collection and analysis plan for Research Question 2, and how it links with the analysis plan for Research Question 1.

Table 4.10: Research Approaches (Questions 1 and 2)

RQ	Enquiry Type	Analysis level	Evidence
1	1 Feedback intention (deductive)	Message and move	Content themes [Pilot codes] (Johnson, 2013)
1	2 Exploratory word use (deductive)	Word	Exploratory word frequency
1	3 Specialized word use (inductive)	Word	Keyword analysis
1	4 Discourse management (inductive)	Between/across message	Rephrasing, revisiting, questioning, and turn taking regularities
2	5 Feedback intention and reception (inductive and deductive)	Message, move and word	TL and examiner perceptions

My analysis of effectiveness involved building in some case study opportunities that engaged a sample of participants in self-reflection about their feedback experiences. This reflection focused on TLs' and examiners' intentions and reactions to particular feedback messages, and allowed me to assume greater confidence in the interpretations that I made during my data analysis.

A key challenge for this research project was to ensure that the data were gathered without influencing the process under investigation. This phenomenon is referred to as a Hawthorne Effect (Festinger & Katz, 1953), and highlights a key problem for social research studies which relates to the issue of establishing a 'closed system where no event outside of it influences events within it' (Cook, 1962, p.120). Although there is a degree of scepticism about whether this effect exists across all contexts (e.g. Wickström & Bendix, 2000), there is some evidence from education research that participants can alter their behaviours when participating in studies (Barnes, 2010). As a result, I felt that it was imperative that I take this issue seriously. In the context of my study, any alteration to examiners' normal professional

behaviour would be an invalid influence on the results awarded to examination candidates, and would potentially impact their subsequent opportunities of access to university or work. This was a risk that I could not take, and forms an aspect of the ethical considerations that I needed to take into account during my research planning phase (see Section 4.4 for more on the ethical issues relating to the study).

I used interviews that employed a Stimulated Recall (SR) technique to carry out a series of individual case studies centring on each TL. These case studies were carried out twice with each TL over a two-year period, meaning that there were a total of six case studies. Each case study focused on specific instances of feedback discourse, and involved 14 examiners. The details of the case study planning are described below in sections 4.3.1.1 and 4.3.1.2.

Case study approaches are influenced by ethnography, and focus on capturing the complexity of a bounded case (e.g. a unit of activity). According to Hamilton & Corbett-Whittier (2013), case studies are useful for investigating the interaction and communication between cases and the outside world. The approach uses a variety of data gathering techniques, and these can support the endeavour of the researcher to triangulate evidence and add legitimacy to their conclusions. My use falls within a descriptive, rather than an exploratory or explanatory frame (Yin, 2013). This is because I seek to gather information on the particular features of feedback discourse, which are themselves informed by the theoretical perspective that steers my data collection approach. According to Merriam (1998), this element of my study can be characterised as being a heuristic psychological case study. This is because I focus on the thinking of one person in situ, to 'bring about the discovery of new meaning… or [confirm] what is known' (Merriam, 1998, p. 30).

My choice of this methodological approach was dictated by the scant research currently available on the nature of feedback-giving. The method also allows a holistic and in-depth investigation of complex phenomena, and so was appropriate because the topic of study (i.e. discourse) is theorised as being inseparable from its context (Yin, 1981, p. 99). As such, I used the method to be able to understand something of the context of feedback-giving; for example, how to better understand the specialised language and concepts that are used in the process of giving feedback. Appropriate methods are required for this type of study because interaction presents a challenge for non-experts (i.e. those outside a specialised professional community) when they try to understand how experts convey information and build meaning through communication with each other. Non-experts have only a limited framework which

they can use to interpret communicated content. Moreover, expert communication involves efficiencies in the form of taken-for-granted elements, which remain tacit as they do not need to be explained to co-participants. The challenge is for non-experts to structure the context so that experts are encouraged to consider and articulate tacit elements of their communication process.

Gomm, Hammersley, & Foster (2000) caution against the ability to generalise from case study data. This criticism is based on the context and sample bound nature of the data gathered. At the same time, there is no a priori reason why the setting and participants in my study are unrepresentative of other examiners working in professional assessment contexts. Although my sample of TLs and examiners was opportunistic, the examiners were expert practitioners carrying out professional activities across a range of subject areas. Their professional status was sanctioned by the highly-regulated OCR Awarding Body that commissioned them, and to that extent there was nothing to suggest that their behaviours would be unrepresentative of other TLs and examiners commissioned by that (or any other regulated) awarding body. This approach is justified by the assumptions that underpin the operating procedures of OCR. These assumptions hold that the practices of examiners are necessarily generalisable to other examiners as there are regulated QA procedures that ensure that examiners' practices are consistent across the organisation.

In the next section I discuss (1) the strengths and weaknesses associated with the use of interview methods, with specific reference to the use of a particular introspective method (Stimulated Recall [SR]); (2) the reasons behind how I constructed my interview and observation schedules; and (3) my approach to analysing the data that resulted from the use of these approaches.

4.3.1 Problematising the Interview and Stimulated Recall (SR) Method

4.3.1.1 Interview Methods: Strengths and Weaknesses

Interview methods have some advantages for use in social science investigation. They employ verbal interaction techniques and can use direct or indirect questions for a number of purposes, including (inductive) data elicitation data or (deductive) hypothesis testing. Cohen and Mannion (1994) note that there is a clear link between interview methods and the types of data that can be collected, with interview data being able to give insights into the core features of participant's thinking. Through careful elicitation, they can also gather insights into knowledge and understandings that are embodied, where a participant may be 'knowing

but incapable of telling' (Alvesson, 2011, p.29), or gather data on privileged opinions, feelings, emotions and experiences (Denscombe, 2010). According to Frankfort-Nachmias and Nachmias (1996), the contextualised nature of interviews is useful, allowing elicitation to take into consideration features of a participant's environment or additional factors that can be used to help data interpretation. This also has potentially positive links to validity, since the interview method allows checks to be made on the accuracy and relevance of data gathered during data collection (Denscombe, 2010, p.192).

At the same time, interview methods have a number of weaknesses that need to be recognised by researchers when designing studies. A number of concerns relate to the potential for subjectivity and bias to undermine the validity of data gathered via interview methods. This means that there is a need to guard against complacent assumptions that any particular interview question necessarily measures a particular phenomenon simply based on the use of apparently transparent question language. A number of balance and control measures can be introduced into the interview development process to aid valid data gathering (Kvale, 1996). One measure is to develop a clear frame to the interview that defines the context of the study for the participants. Such a process can minimise the chance that a researcher imposes their own views of reality, (or their presumptions about a participant's views of reality), on the interpretations they place on participant's responses (Foddy, 1993).

Interview methods also tend to use a limited sample of participants, potentially undermining the ability to make generalisations from the gathered data. This means that there needs to be a clear rationale for sample selection so that the limitations of any particular selection decision are recognised (LeCompte & Preissle, 1993).

4.3.1.2 SR Methods: Strengths and Weaknesses

Stimulated Recall is an introspective interview approach that prompts a participant to recall thoughts from an earlier task. It helps to overcome the weakness that is associated with interviews where participants are required to rely heavily on memory structures when reporting information about thought processes. SR includes the use of tangible prompts (e.g. verbal, photographic or video materials) as a reminder of past behaviour that stimulates recall of mental processes in operation during the task.

SR is an approach that some argue helps to structure a participant's responses (Gass & Mackey, 2000). In similar projects SR approaches have been used to gain insights into examiner thinking whilst seeking not to unduly influence the mental processes being

observed (Crisp & Johnson, 2007; Johnson & Black, 2012a), and so such methods have been used to help mitigate the potential effects of the observer's paradox (Labov, 1972, p.61). Lyle (2003) notes that SR methods have been used extensively in educational research. The methods facilitate introspection procedures through focusing on prompts from a 'video sequence or some other form of visual recall' (Fox-Turnbull, 2009, p. 204) to stimulate recall of the cognitive activity that participants experienced concurrent with an observed episode of behaviour. This relationship between objects and memory is also a strong feature for studies that focus on the links between material and culture (e.g. Kirk & Sellen, 2010; Morgan & Pritchard, 2005).

Considerations about the validity of recall are the largest area of concern for critics of SR methods. As Calderhead (1981) points out, the validity of participants' 'recalled thoughts cannot be rigorously checked' (Calderhead, 1981, p. 215), but the researcher can look for evidence of internal consistency (or lack of contradiction) in the data elicited. Notable criticisms of the use of SR include concerns that the process actually leads to participants supplementing their incomplete memories (Lee, Landkin, & Carter, 1992); that the participant is reacting to what they are seeing during the playback (Tjeerdsma, 1997); that the recall might not represent the conscious or unconscious cognitions taking place at the time of the episode (Wilcox & Trudel, 1998); that recall might produce a 'new view' to the participant which is 'subject to the luxury of meta-analysis and reflection' (Yinger, 1986); or that 'recall decay' might lead to there being a difference between the 'recall of' and 'reflection on' an event (Gass, 2001).

There are claims that participants' retrospective reports have been found to be as accurate as concurrent reports when obtained immediately after a task (Norris, 1990) and that recall accuracy can be in the region of 95% for up to 48 hours after an event (Bloom, 1953). SR has advantages over other introspective methods (such as think aloud) as it not reliant on memory and it does not interfere with the task or with the participant's short term memory when carrying out a task. As a result, proponents of SR claim that it can be a highly valid representation of a participant's thinking so long as a number of concerns are taken into consideration at the research design stage (Gass and Mackey, 2000; Lyle, 2003). Concerns around the validity of the method have led to the development of guidelines for best practice. In their review of how to use SR techniques, Gass and Mackey (2000) insist that for the recall method to be reliable and valid it is important that there is minimal time delay between event and recall. This is reiterated in the findings of Cassaday, Bloomfield and Hayward (2002),

who argue that material is best remembered if circumstances of learning and retrieval are the same, and if the presence of some of the cues that were around at the information encoding stage are used as reminders.

4.3.2 My SR Interview Method

I chose to use interviews because they were appropriate for the needs of my particular research questions – which were to identify the characteristics of TL feedback communication, and to explore participants' perspectives on the features that were effective in terms of aligning their understandings. As a result I wanted to gather information that was privileged to the participants as a consequence of their professional role, and interview data is able to give insights into the core, embodied, or privileged features of participant's thinking (Alvesson, 2011; Cohen and Mannion, 1994; Denscombe, 2010).

Since I also wanted to explore contextualised interaction, I considered interview methods that accommodate SR techniques to be an appropriate data collection method because they can take into consideration features of a participant's environment (Frankfort-Nachmias and Nachmias, 1996).

I was aware of a number of concerns that relate to the potential for subjectivity and bias to undermine the validity of data gathered via interview methods. To guard against complacent assumptions that any particular interview question necessarily measures a particular phenomenon I needed to build a number of balance and control measures into the interview development process. For this I looked to Kvale (1996), who outlines a number of steps that influence the quality of interviews. One such step was to develop a clear frame to the interview that defines the context of the study for the participants. My use of follow up member checks, or 're-interviews' (Kvale, 1996, p. 190), also allowed me to check that the interpretations that I was making from the interview data were triangulated with the interpretations of the participants.

Interview methods tend to use a limited sample of participants, potentially undermining the ability to make generalisations from the gathered data. For this project, it was legitimate to use non-probability purposive sampling to include participants who had experiences that were in common with each other, but that also differed in some particular ways. For example, involving TLs from different subject groups capitalised on their commonly shared examining experiences, whilst also allowing for possible differences resulting from the subject domains within which they worked. Whilst such sampling inevitably limits the generalizability of the

research outcomes, it was tailored to the needs of the research design to elicit specifically positioned perspectives.

My approach to interview design in this study took into consideration a number of theoretical and logistic factors that were prompted by temporal and geographical issues, including my relationship with the examiners. My position as a researcher attached to the organisation that contracts work to the examiners potentially left examiners feeling vulnerable and reluctant to participate. As a result, the initial contacts made with examiners were constructed to create some distance between my professional role and my current study work by highlighting the institutional link between the study and the University of Cambridge Faculty of Education. I also chose to spend time building up professional trust through personal, face-to-face interactions with the TLs in the run up to the study. This allowed them to be able to ask questions about the aims and intentions of the study.

The situated nature of the study suggested that an important aspect of the validity of using interview methods linked with its ability to elicit data as close as possible to participants' involvement with actual feedback processes. There were obvious methodological and logistic difficulties in gathering concurrent data during shared communication, so examiner interviews were carried out in two stages to cope with this challenge. Earlier evidence suggested that examiner feedback was largely hierarchic in character (Johnson, 2013), so I felt that this legitimised a data gathering approach that initially focused on the TL at the moment of feedback generation. In such a case, interviews with TLs gained insights into the decisions that shaped the character of feedback, in effect giving TLs the opportunity to share their perspectives on their feedback crafting decisions.TL interviews were thus the first interview stage.

I wanted to elicit examiner's perceptions around their received feedback. In a past project, this has been accomplished through the use of telephone interviews with examiners (Johnson & Black, 2012b). Bearing in mind Kvale's (1996) concerns about the affordances of face-to-face interview methods I sought to carry out face-to-face interviews, but due to the wide geographical spread of examiners this was not always possible. In cases where the travel time to an examiner would take more than three hours, or where their accessibility was only late in the evening, I used a telephone interview method (Table 4.11).

I carried out all of the interviews as close as possible to the time of feedback receipt. This helped to ensure that the memory trace of their initial perspective on the feedback was still

accessible. To facilitate the interview process, I provided all examiners with an advanced copy of the feedback that was to be the focus of the interview.

Table 4.11: TL and Examiner Demographic Data

Chemistry					Economics					Geography				
TL1: *Male; Age: 50s; Examiner for 20+ years; TL for 8 years; a retired teacher of Chemistry from a Comprehensive School; still continues to tutor students privately*					TL2: *Male; Age: 20s; Examiner for 5 years; TL for 2 years; currently a teacher of Economics in a Comprehensive School*					TL3: *Female; Age 50s; Examiner for 20+ years; TL for 5 years; a retired teacher of Geography from a Comprehensive Sixth Form College*				
Examiner	♀♂	Exp[1]	Fam[2]	Int[3]	Examiner	♀♂	Exp	Fam	Int	Examiner	♀♂	Exp	Fam	Int
1	F	*	*		8	M	*	*	†	14	M			
2	M	*	*	†	9	F				15	M	*	*	
3	F				10	F	*	*		16	F	*	*	
4	F	*	*		11	M			†	17	M			
5	M			†	12	M	*		†	18	M			*
6	M	*	*		13	F				28	M	*	*	
7	F				24	M			†	29	F			*
19	F			*	25	M	*	*	†	30	M			
20	M	*	*		26	M	*	*	†					
21	M			†	27	M	*	*						
22	M	*												
23	F	*	*	*										

[1] Experience (the examiner has examined in a previous marking session); [2] Familiar to the TL (worked together previously); [3] Interviewed examiner (*Face-to-face/†Virtual)

Finally, to satisfy demands of construct validity, I wanted to ensure that the interview respondents could define the topic about which they were responding in the same way. As a result, the responsibility fell on me to stipulate and define as clearly as possible the dimensions of focus. Interviewer variance is a concern for critics of interview methodology, and this is an issue that relates to the social dynamics inherent to the interview setting. One way of reducing this issue was through imposing a standardised schedule that helped to structure the interview interaction. It has been recognised that by manipulating the conditions of an interview it is possible to influence the point at which the researcher imposes structure on information to produce data (Brown & Dowling, 1998). These manipulations allow interviews to be categorised into three groups - structured, semi-structured, and unstructured - each differing in the degree of researcher or participant control over the issue agenda (Corbin & Morse, 2003). I chose to use semi-structured interviews to gather data about TLs' and examiners' perspectives because this type of interview allowed me to retain some control over the interview structure whilst also allowing the participant to shape the interview outcomes with their own particular perspective. Using pre-specified questions, it is assumed that participants' responses will only be partially elicited through unmediated interview

questions, with social interaction helping to elicit valid responses through additional probing and the creation of interpretive space (May, 2011). In this way, a flexible semi-structured interview approach enables spaces to be created for divergent interpretations to be explored. Semi-structured interviews also enable the participant to move beyond mere description of the phenomenon of interest to potentially noticing new relations through reflection during the interview process (Kvale, 1996). I therefore chose to develop a semi-structured interview schedule to allow scope to probe beyond the interview participants' answers and to enable me to search for clarification and elaboration. I felt that this was a valid approach because, according to May (2011), it allows 'an opening up of the interview method to an understanding of how interviewees generate and deploy meaning in social life' (May, 2011, p.135). In the next section I describe how I carried out my interviews and how the data was analysed.

The semi-structured interview schedule that I used steered the TLs and examiners to focus on the global- and sentence-level content elements of communication that coalesced around the emerging ASCDA framework. These elements of analysis were also informed by the research that I gathered in my literature review that looked at feedback communication in other contexts (these elements are shown in Appendix J). The integration of the findings from literature allowed my analysis to be informed by theory, helping me to establish a methodical stance with some distance between myself and what was said by the interview participants (Kvale, 1996, p. 201). These elements then formed the basis for exploring whether intention and reception mapped closely onto each other, and whether there were any variations between how these communication elements successfully supported aligned understanding. This function was reiterated by the way that the data from both participants was organised so that it was presented alongside each other (Appendix K). The schedule had a sharp focus on the feedback giver's intended objective and predicted impact, with this being contrasted with the recipient's perception of what the message achieved (actual impact). To facilitate this, questions about the feedback messages were probed through different units of analysis: at the levels of message structure, sentence units, and paragraph (to allow links between sentences to be accommodated).

Each TL observation session had two elements: the initial observation of the TL crafting their feedback; and then the SR interview session that immediately followed this. During the initial observation, which I carried out in a room next to that used by TL, I used software that allowed me to see the TL at work through a webcam as well as the contents of their computer

screen (see Figure 4.3 below). At this time I was able to code each feedback using a framework that broke each message up into its constituent sentences, in order that I could focus on the elements that I had identified as being of interest in my theoretical framework (Appendix J). Using this framework I organised each message into a number of sentences and considered *what information was being dealt* with in each line (e.g. feedback or technical content); *how this information was being offered* (e.g. as a statement, a location, a description, a clarification or a directive); whether any *bridging or referencing strategy* was being employed (e.g. internal, external etc.); and whether any *stylistic/interpersonal influences* were evident (e.g. upgrading and downgrading emphasis).

I then used this gathered data as the basis for my SR interview with the TL, and then as a basis for comparison with the perspectives of examiners when I subsequently asked them about the same feedback message. I was able to probe the element of feedback *impact* through my TL and examiner interviews by asking the participants whether: (1) there was an intended (or perceived, in the case of the examiner) objective for the feedback message; and (2) whether the feedback helped the examiner to have learned something. The interview schedule for these sessions is presented in Appendix K and shows how the questions probe the participants to consider the feedback at both global (i.e. overall intention and effect) and constituent (i.e. sentence and word) levels. For example, the examiners were asked at a global level '*If you saw a performance like this again do you think that they would be in a position to know how it should be marked as a result of this feedback message?*' and at a constituent level '*Is there an element of the feedback message that you find most useful and why?*' To analyse the data I was able to compare the TLs intentions with those reported as being perceived by the examiner for each message.

Below are screenshot examples of a section of video playback that was part of a SR interview (Figure 4.3). The first screenshot shows the TL attending to the examiner's marked examination script at 5 minutes 28 seconds into the session, with this followed by the crafting of feedback at 6 minutes 24 seconds.

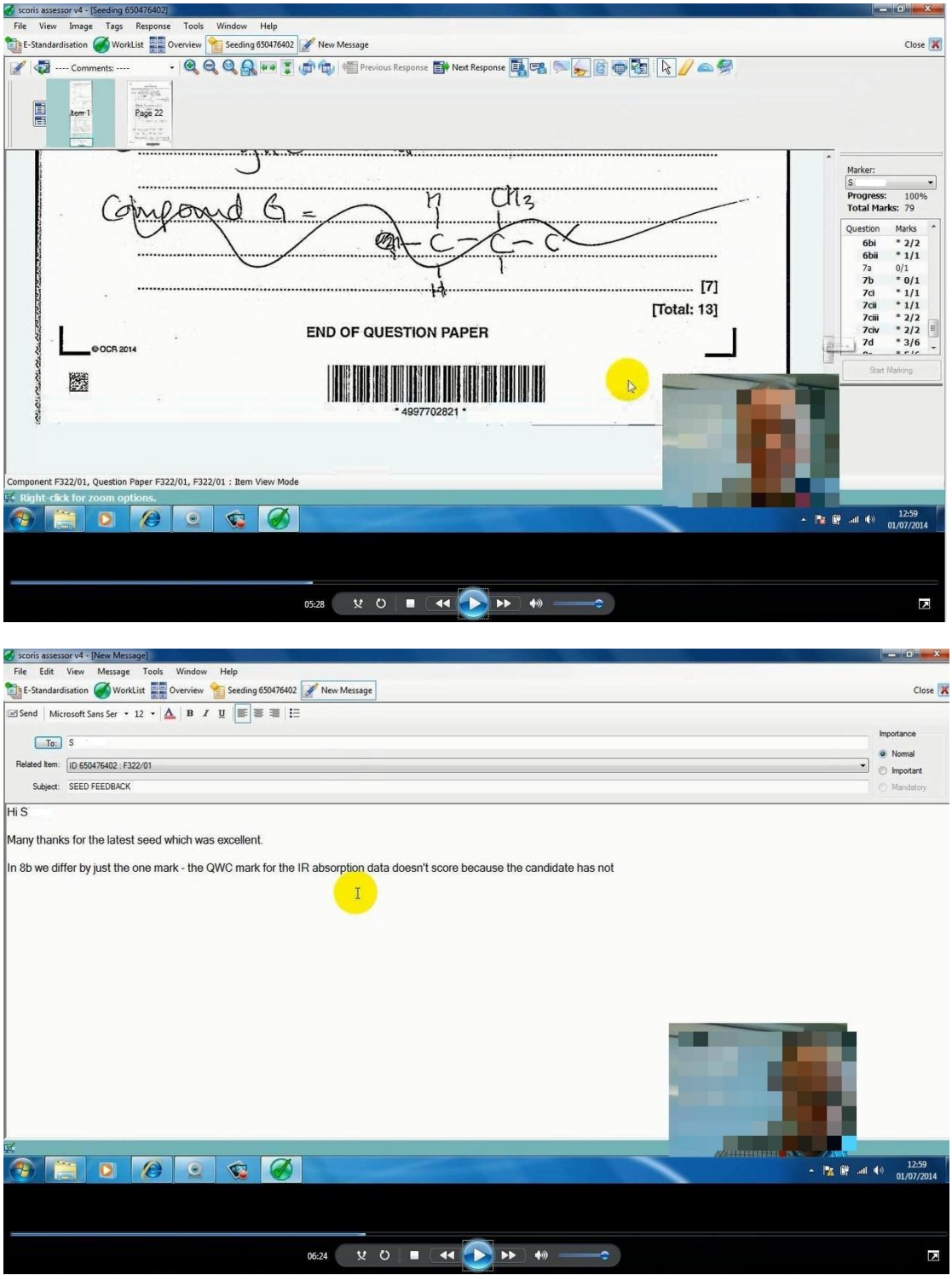

Figure 4.3: Screenshots of TL Video Playback

The examiner interviews were then transcribed by a transcription agency. I quality assured the transcripts by listening to each recording and checking that the transcription matched the interview recording. For ease of interpretation, in this study I present interview data in italicised and indented font. Where a participant quotes another person, this is shown in

inverted commas. Additional contextual information that is added to aid the reading of the transcript is indicated by square brackets.

> *I'll open up [script] 2443, 1a. Oh no, that's not okay, well maybe it's the other question, or it's not there, right, which one is it? And you know ultimately it's just a slip of the finger but it then means you probably spent maybe five minutes or so going through your other papers, and by that point you are just thinking 'which is the one?'*
>
> *[Y1 E18 SR interview 104]*

The next stage of the process was to analyse the transcribed data. According to Kvale (1996), analysis involves data reduction prior to interpretation (which involves data expansion). My approach to analysis involved taking each subject as a separate case and coding the TL transcript using MAXqda textual analysis software (VERBI Software – Consult – Sozialforschung GmbH, 2013).

This coding involved tagging each reference to an item in the interview transcript. For each item, I was able to record what the TL's expressed intention behind the message was, and how they anticipated that the examiner would receive it and act on it. These records then allowed me to look for evidence of alignment between intention and interpretation in the interview transcripts. In keeping with concerns expressed by Kvale (1996), these interviews were transcribed verbatim so that qualitative insights could be gleaned.

The next stage of the analysis was to tag each question item in each examiner interview and to map these to the TL transcript. For each item the examiner was asked to explain what the effect of the feedback was on them, what action they carried out, and whether they could predict the TL's intention behind the feedback. This analysis helped me to explore the existence of *impact*, as this was one of the dimensions of interest in my ASCDA theoretical framework. Figure 4.4 shows an example of how an instance of an examiner's message perception (highlighted in the top right window) compares with the TL's intention behind the message (bottom right window).

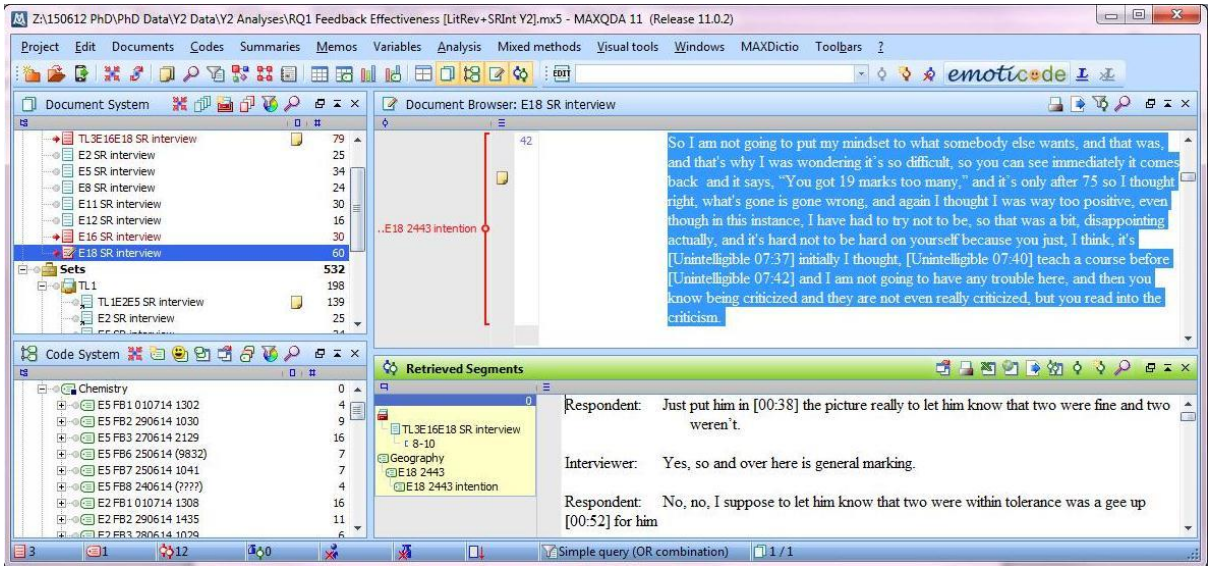

Figure 4.4: Analysing Intention and Reception [Screenshot of MAXqda Output]

This coding structure enabled me to carry out two analytical functions. Firstly, I was able to map the intentions and perceptions of each participant examiner to those of their TL for each message. This allowed me to consider whether there was alignment or misalignment between the participants on each item. Secondly, where alignment was found I was able to consider the features that appeared to contribute to it.

One concern about the use of interview data is that it relies on self-report information and those who subscribe to a Conversation Analysis or ethnomethodological approach (which in part also informs my study) would argue that evidence of the actual demonstration of alignment should also be considered. To deal with this concern I also wanted to analyse instances of feedback communication where the participants achieve resolution through their interaction. My coding framework affords the opportunity to consider this issue by allowing me to identify a subsample of data where there is evidence of resolution. According to LeCompte & Preissle (1993, p. 82), when discussing the strengths of qualitative analysis, it is valid to sample data from a broader population if it helps to illuminate a particular phenomenon. They also argue that the study of subsamples using qualitative analysis can reveal things that would not otherwise be recognised due to a 'lack of statistical power' in the context of a conventional quantitative analysis. Discussing ethnographic study design LeCompte & Preissle (1993, p. 75) argue that a 'unique-case selection' is legitimate when researchers want to examine some dimension where a close analysis highlights the processes

that operate within a select group, and therefore this form of analysis is not based on statistical relations.

In order to look more closely at the conditions that might lead to resolution in discourse, I adopted a strategy to identify a sub sample of cases where resolution was evidenced (through my TCA coding approach). In keeping with the CA-approach that partly informed my study, this allowed the participants' own indications of resolution to steer my analysis and to indicate where agreement was displayed.

I chose a targeted sample of 55 episodes of interest (from 329 instances of feedback data where there was evidence of two-way interaction; for more on this see Chapter 6). To carry out this form of analysis I identified the codes that interconnected with the 'resolution' code across a number of cases. This allowed my analysis to focus on the qualitative characteristics that appear to be present across the cases. This targeted approach is in keeping with similar studies where elements of educational discourse of interest are taken as exemplars for the purposes of discussion and generalisation (e.g. Mercer, 2004). It is also in keeping with the ethnographic basis that some argue needs to underpin sociocultural studies. Gee & Green (1998) argue that an ethnographic basis is important because they claim that interpretations need to be ethnographically grounded whilst affording a view of the complex patterns that are constructed across different interactions (Gee & Green, 1998, p. 158).

Although I have alluded in passing to some of the ethical considerations that I encountered during this research, in the next section I identify a number of specific ethical issues and describe how I dealt with them.

4.4 Ethics

The high stakes assessment setting of my project, along with my own professional position within an organisation that had links to the OCR Awarding Body, meant that I needed to consider a number of ethical issues when planning and carrying out my research project. According to Hammersley & Traianou (2012), these issues would have implications for the methods that I chose to employ when collecting data, and for the way that I conducted myself as a researcher in my interactions with participants in the research project.

With regards to data collection, the requirement to gain voluntary informed consent from research participants before research gets underway is an established norm set out in guidelines laid out by the British Educational Research Association (British Educational

Research Association, 2012). In order to give informed consent, my participants needed to understand the nature of the project and how the data which they helped to generate were to be used. The first stage of the informed consent process involved me informing the participants about the purpose, scope and details of the project. For each of the two phases of the project I produced an information sheet with details of the aims of the project and the nature of the participants' anticipated involvement (Appendix L). This disclosure helped to mitigate a potential ethical threat that participants might be unaware of the actual aims of a particular research project (Hammersley & Traianou, 2012).

Another key element of informed consent was that agreement to participate should not be carried out under duress or pressure, and that participants should be allowed to withdraw from involvement in the research project at any point. This concept is termed voluntarism, and Frankfort-Nachmias & Nachmias (1996) suggest that this may partly be established through the researcher forging an egalitarian relationship with the participants. I sought to establish an egalitarian approach by ensuring that the participants' contribution to the project was recompensed at a rate that was proportionate to the work involved. To do this, I paid the TLs and examiners at a level that matched their usual contracted pay rates when employed by OCR to carry out marking work.

TL and examiner feedback data was collected through either accessing the TLs' *scoris*® *assessor* accounts (where their email feedback communication history was stored), or through gathering recordings of their feedback telephone conversations. Prior to collecting any feedback communication (email or telephone), the examiners had to indicate that they consented to letting their TL share their feedback communication with me (Appendix F). Of the 34 examiners who were approached to participate in the research project, four declined. This meant that any data that directly related to these examiners was not captured. To access the *scoris*® *assessor* accounts, the TL supplied me with their account login details. For telephone recordings, the TL asked the examiner at the start of each feedback conversation whether they were happy for their conversation to be audio recorded. No examiners refused this request.

Once collected, the study data was stored on a secure, password protected data system within the Cambridge Assessment organisation. This ensured that the data was held in a controlled environment that was protected by an organisation-wide policy to maintain security standards (Cambridge Assessment, 2015). This action ensured that confidentiality was prioritised

(Social Research Association, 2003, p. 38), and minimised the ethical risk that sensitive information might become publicly available that might damage the reputations of the participants (Hammersley & Traianou, 2012).

As mentioned above, the relationship that I established with the project participants had ethical dimensions. For example, participant perceptions about the amount of freedom that they had when taking part in a study could be influenced as much by the data collection methods adopted by the researcher as by the information that was conveyed to the participant prior to their consent to participate. Cassell (1980) notes that the locus of control in a research context is a dimension of the relationship that develops between a researcher and a participant. By manipulating the conditions of an interview, it is possible to influence the point at which the researcher imposes structure on information to produce data (Brown & Dowling, 1998). Issues of power relationships in interviews are an important consideration. This is because, for example, interview processes can disempower the participant through the way that the researcher sets the agenda for the project (Limerick, Burgess-Limerick, & Grace, 1996). The ethical problems of suggestibility have been pursued in studies of research with young children (Garven, Wood, & Malpass, 2000; Garven, Wood, Malpass, & Shaw III, 1998), but this is also an area of concern for research involving adult participants (Lindberg, Keiffer, & Thomas, 2000). In my study, I used semi-structured interview procedures as much as possible so as to reduce the potential ethical threat of leading the participants to give socially-influenced responses.

Collecting data in a high stakes professional assessment context also raised considerable ethical issues. At the first level, there were concerns that approaching participants through their employer may have an implied coercive element. I was concerned that participants may feel compelled to participate if I approached them via my link with the Cambridge Assessment organisation (which is the parent organisation of OCR, which normally commissioned their work). As a consequence, my invitation to the examiners distanced my connection to OCR and foregrounded the study link with the University of Cambridge Faculty of Education (Appendices F and L). At the same time, whilst making clear my affiliation with the Faculty of Education, I also made it clear that my work was funded by Cambridge Assessment. This allowed me to highlight the link between how participation in the project could contribute to improving future TL and examiner practice.

At a second level I had concerns that information about the participants' performance collected during the project could adversely affect their professional status. According to the Social Research Association, harm to subjects may arise from undue stress through loss of self-esteem (Social Research Association, 2003, p. 35). As a consequence, it was important to maintain confidentiality of identity in both the way that the examiner data was stored, and in the way that the study outcomes were reported.

Finally, the high stakes assessment context raised particular concerns that the research study could influence people who were not directly involved in the research study (i.e. the students who were being assessed during the research study data collection). This tension is a feature of research into examinations, which has traditionally sought to assure that research studies do not interfere with live marking processes and avoid the problem of the observer paradox (c.f. Labov, 1972; Landsberger, 1958). In this way, the adoption of specific research methods has a direct ethical bearing. As a consequence, I chose to use observation and interview methods that minimised intrusion into examiner practice as much as possible.

4.5 Summary of My Methodological Approach

In this chapter I describe how the methods that I used for my pilot study influenced the methods that I adopted for the main study. I also outline how I constructed these methods based on a sociocultural perspective that helps me to address my two research questions:

1. What is examiner feedback? and;

2. What is effective about such examiner feedback?

For the first research question, and reflecting the Vygotskian heritage of sociocultural theory which holds that individual development is mediated through cultural tools which are developed through social interaction (Vygotsky, 1978), my analyses needed to evidence the interpersonal and contextualised nature of communication. To do this I explain in detail my ASCDA methodology that draws on tools that are commonly associated with Discourse [Thematic Content] Analysis, Conversation Analysis, and Corpus Linguistics approaches. This methodology enables analyses with both qualitative and quantitative characteristics to explore both the contextualised and general dimensions of feedback communication according to the themes of *Content, Time, Joint Intellectual Action,* and, *Impact*.

For the second research question I employ case study, semi-structured interview, and observation methods to explore how participants conceptualised feedback effectiveness. I am then able to relate these outcomes to theoretical evidence of effectiveness (from literature review), that I have previously organised into a framework (Integrated Analytic Framework) in Chapter 2. Finally, to overcome some of the limitations of self-report data I analyses some case studies where examiners demonstrate the construction of alignment through feedback discourse.

5. Results: What are the Characteristics of Examiner Feedback?

In the first section of this chapter I describe the characteristics of the feedback data that I gathered during the research project. In the second section I analyse these data using the ASCDA methodology that I described in Chapter 4.

Before looking in depth at the characteristics of the feedback data at a qualitative level (using the ASCDA methodology) it is useful to consider the general nature of the data. This allows any discernible patterns that might be noticeable through a quantitative analysis to steer the qualitative enquiry. For example, differences in the quantity of feedback information being communicated may be indicative of the trouble sources (i.e. areas of weak between-examiner common ground) that necessitate feedback activity. These trouble sources may pertain to examiner characteristics (such as experience and familiarity), or the stability of the area of knowledge and the mark scheme type being discussed (see Section 5.1 for more on this).

As I outlined in the previous chapter, feedback data gathering took place over two examination marking sessions, and included a total of 991 feedback messages. These messages covered all of the communication that took place between TLs and examiners, and included messages from an examiner to a TL. Except for where I indicate otherwise (e.g. in cases where I look exclusively at TL word use) my analyses include all feedback messages that were communicated between a TL and an examiner.

There were on average 33 messages communicated between each TL and each examiner, with this ranging from 21 messages (Chemistry) to 47 messages (Geography). The majority of these messages were communicated via email. Table 5.1 shows the distribution of the feedback messages and word count across subject areas, examiner experience levels, examiner familiarity levels, feedback communication modes, and across the body of the discourse (in terms of quartiles).

Table 5.1: Feedback Quantity

	Messages	Word Count
Chemistry	248	29,571
Economics	371	38,489
Geography	372	110,075
Experienced	524	76,968
New	467	101,167
Familiar	442	70,045
Unfamiliar	549	108,090
Telephone	15	19,572
Email	976	158,561
Quartile 1	248	59,688
Quartile 2	248	50,182
Quartile 3	248	32,448
Quartile 4	247	35,815
Total	991	178,135

To get an idea of the characteristics of the feedback message data I wanted to compare how much feedback, and by inference, the amount of content that was being communicated in the feedback messages. This analysis would allow me to investigate whether feedback was responsive to particular trouble sources (e.g. weak common ground pertaining to examiner experience and familiarity, or the stability of the area of knowledge and the mark scheme type being discussed). I used 'word count' as a proxy for the amount of content that was contained in the feedback messages that passed between TLs and examiners because this was a useful indicator of feedback volume.

Building on the literature and the theoretical framework that I have already outlined I was interested in whether there were any initial indications that examiner characteristics (such as their prior examining experience or familiarity with their TL), or the nature of what they focused their interaction on (such as item type) would influence the amount of feedback being communicated.

5.1 Feedback Content Volume (Word Count Analysis)

My working hypothesis was that the amount of feedback content could be an indicator of the amount and nature of learning discourse being communicated, which could lead to further exploration through qualitative methods through my ASCDA approaches.

This hypothesis is based on my theoretical framework, which anticipates that feedback interaction would act as a repair mechanism and centre on a trouble source. I have outlined earlier how there tends to be more examiner disagreement around levels-based marked items, suggesting that these types of items are a weak foundation for building shared

understandings. As a result, I was interested in whether there would be more feedback that focuses on levels-based marked items (which were more prevalent in the Geography question papers compared with the Economics or Chemistry question papers[15]).

My hypothesis also proposes that TL and examiners' common ground establishment, which includes work related to the construction of an Intermental Development Zone (IDZ), is a component of learning discourse, and that this work would be most pronounced at the earliest phase of feedback interaction. As a result, I was interested in whether there would be more feedback at the earliest marking stages compared with the later marking stages.

In addition, since I claim that feedback is a tool for inducting professionals into a community (c.f. Basturkmen et al., 2014), through building common ground via establishing shared language and ways of thinking, I was interested in whether there were any indications that experienced examiners would require less induction than new examiners. This coheres with the implications that I draw from my Integrated Analytical Framework (IAF) that discourse that seeks to support the construction of learning has a participative dimension and an active recipient role (and so would include more content).

And finally, I was interested in whether examiner familiarity with a TL could also be a feature that influenced the construction of the IDZ. Regardless of an examiner's experience level it is possible that examiners who have worked together will have accrued more information about each other and will understand each other's needs more than those who are new to each other. This also coheres with the implications that I draw from my IAF that highlights how effective feedback would contain some implicit elements that centre on relationship management.

To explore these four areas I set out to analyse the mean word count for feedback messages that related to: (1) subject area (i.e. being a proxy for marking complexity. The Geography paper contained the highest proportion of subjective items and these are considered to be the most complex item type and result in lower levels of examiner agreement); (2) time (i.e. through splitting the discourse for each TL into four equivalent chronological sections; (3) examiner experience level (i.e. whether they had examined previously or not), and (4) TL-examiner familiarity (i.e. had the participants previously worked together, or not).

[15] Geography included nine Levels-based marked items in the paper for each year (82% of each paper) compared with Economics that had one Levels-based marked item in the paper for each year (8% of each paper) and with Chemistry that had no Levels-based marked item in the paper for each year.

I chose to use word count as a way of comparing across my areas of focus because it is a simple measure of the amount of information being communicated. I could have approached this analysis through measuring the number of messages that were communicated, but this would have been less transparent as a unit of analysis as all messages would not be equal since some message would contain more information than others.

To choose an appropriate method for analysing any differences in the mean word count of feedback messages for the different subject groups I needed to check for assumptions of normality. A Shapiro-Wilk normality test (Table 5.2) showed that the data was not normally distributed ($W = .373$, $p = <.001$), meaning that I needed to use a non-parametric test to compare the group word count means for my areas of focus.

Table 5.2: Shapiro-Wilk Normality Test for the Word Count Data

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
FBwordcount	.298	991	.000	.373	991	.000

a. Lilliefors Significance Correction

5.1.1.1 Feedback Volume and Subject (as proxy for item type)

To compare the word count data across the feedback messages across the different subject teams (Chemistry [C], Economics [E], Geography) [G]) I chose to use the Kruskal Wallis H nonparametric test. This test can be used for comparing the means of more than two samples (e.g. the word count differences across three subjects in my sample or across the four quartiles of feedback time) (Field & Hole, 2003). The data satisfied the other assumptions for the test to be used (i.e. my dependent variable was continuous, there was independence of observations, and the distribution for each group of the independent variable had the same shape).

Table 5.3: Kruskal Wallis H Nonparametric Test (Subject Word Count)

Ranks

	Subject	N	Mean Rank
FBwordcount	C	248	403.76
	E	371	373.98
	G	372	679.19
	Total	991	

Test Statistics[a,b]

	FBwordcount
Chi-Square	245.570
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: Subject

This analysis (Table 5.3) showed that there was a significant difference in the mean word count across the three groups (H = 245.57, p = <.001). An effect size calculation showed this to be a small effect (η^2 = 0.25) (Cohen, 1969).

To establish the location of this significance finding and to identify which subjects accounted for the word count differences, I carried out Dunn's pairwise post hoc tests for the three pairs of groups (i.e. C v E/G; E v C/G; and, G v C/E). There was very strong evidence (p < 0.001, adjusted using the Bonferroni correction) of a difference between the Geography feedback word count and for the word count for the other two subjects. The median word count for Geography messages was 271, compared with 67 (Chemistry), and 64 (Economics).

This analysis provides limited support for the hypothesis that there was more feedback communication for question papers that included levels-based items than for other papers. This also suggests that feedback was a repair mechanism around a trouble source, with levels-based items presenting examiners with a challenge to establish a common shared understanding of 'the correct mark' through their interactions.

5.1.1.2 Feedback Volume and Time

Next I analysed the mean word count according to the feedback quartile (Q1-Q4). In order to carry out this analysis I split the total feedback corpus into four equal sections based on the chronology of interaction for each TL. Again, for the reasons outlined in the last section I chose to use the Kruskal Wallis H nonparametric test to analyse the number of words that were contained in these four sections.

Table 5.4: Kruskal Wallis H Nonparametric Test (Quartile Word Count)

Ranks

	Quartile 1-4	N	Mean Rank
FBwordcount	1st Quarter	248	602.71
	2nd Quarter	248	505.28
	3rd Quarter	248	451.06
	4th Quarter	247	424.66
	Total	991	

Test Statistics[a,b]

	FBwordcount
Chi-Square	56.195
df	3
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: Quartile 1-4

This analysis (Table 5.4) showed that there was a significant difference in the mean word count across the three groups (H = 56.20, p = <.001). An effect size calculation showed this to be a very small effect (η^2 = 0.06).

To establish the location of this significance finding I carried out Dunn's pairwise post hoc tests for the four pairs of groups (i.e. Q1 v Q2/Q3/Q4; Q2 v Q1/Q3/Q4; Q3 v Q1/Q2/Q4; and, Q4 v Q1/Q2/Q3). There was very strong evidence ($p < 0.001$, adjusted using the Bonferroni correction) of a difference between the 1st Quartile and the word count for the other 3 Quartiles. The median word count for messages in Q1 was 176, compared with 88 (Q2), 76 (Q3), and 74 (Q4).

This analysis provides very limited support for the hypothesis that there was more feedback communication at the earliest phase of interaction. This also suggests that at this initial stage the TLs and examiners were using feedback interactions to establish common ground.

5.1.1.3 Feedback Volume and Examiner Experience

In the next analysis I wanted to analyse the mean word count according to examiner experience. As this data category was dichotomous I chose to use a Mann-Whitney U nonparametric test. This test can be used for comparing the means of two samples that are not normally distributed (e.g. the word count of feedback messages between TLs and new and experienced examiners, or between TLs and familiar and unfamiliar examiners) (Frankfort-Nachmias & Nachmias (1996). The data satisfied the other assumptions for the test to be used (i.e. my dependent variable was continuous, there was independence of observations, and the distribution for each group of the independent variable had the same shape).

My analysis of word count according to examiner experience (Table 5.5) showed that there was a significant difference in the mean word count across the experienced and new examiner groups (U = 91596.500, p = <.001, r = 0.22). Feedback communication between TLs and

new examiners contained more words (*Mdn* = 116 words) than the communication between TLs and experienced examiners (*Mdn* = 75 words). An effect size calculation showed this to be a small effect (*r* = 0.22).

Table 5.5: Mann-Whitney U Nonparametric Test (Experienced/New Examiner Word Count)

Ranks

	Experienced Examiner	N	Mean Rank	Sum of Ranks
FBwordcount	Yes	524	437.30	229146.50
	No	467	561.86	262389.50
	Total	991		

Test Statistics[a]

	FBwordcount
Mann-Whitney U	91596.500
Wilcoxon W	229146.500
Z	-6.839
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Experienced Examiner

This analysis provides very limited support for the hypothesis that there was more feedback communication between TLs and new examiners than between TLs and experienced examiners. This also suggests that TLs and new examiners were using feedback interactions to establish common ground in ways that TLs and experienced examiners did not.

5.1.1.4 Feedback Volume and Examiner Familiarity

In the final analysis I wanted to consider the mean word count according to examiner familiarity. Again, as this data category was dichotomous I chose to use a Mann-Whitney U nonparametric test. This analysis (Table 5.6) showed that there was a significant difference in the mean word count across the unfamiliar and the familiar examiner groups (*U* = 100953.000, *p* = <.001, *r* = 0.14). Feedback communication between TLs and unfamiliar contained more words (*Mdn* = 101 words) than the communication between TLs and familiar examiners (*Mdn* = 80 words). An effect size calculation showed this to be a smaller effect compared with that of examiner experience (*r* = 0.14).

Table 5.6: Mann-Whitney U Nonparametric Test (Unfamiliar/Familiar Examiner Word Count)

Ranks

	Prior Professional Relationship	N	Mean Rank	Sum of Ranks
FBwordcount	Yes	442	449.90	198856.00
	No	549	533.11	292680.00
	Total	991		

Test Statistics[a]

	FBwordcount
Mann-Whitney U	100953.000
Wilcoxon W	198856.000
Z	-4.550
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Prior Professional Relationship

This analysis provides very limited support for the hypothesis that there was more feedback communication between TLs and unfamiliar examiners than between TLs and familiar

examiners (but to less of an extent that the difference between TLs and new and experienced examiners). This also suggests that TLs and unfamiliar examiners were using feedback interactions to establish common ground in ways that TLs and familiar examiners did not.

To summarise this section, there is some limited support for the hypothesis that TL and examiner feedback interaction content is focused on the establishment of common ground around the trouble source of levels-based items. There is also some limited support for the hypothesis that the common ground between TLs and new and unfamiliar examiners is weak and that feedback is a mechanism for them to establish shared understandings of marking practice.

These initial and admittedly superficial analyses suggest that there may be some patterns in the feedback interactions and that these may link to how TLs and examiners establish common ground with each other. In the next section I supplement these limited analyses by qualitative approaches (using ASCDA methods) to explore the nature of feedback in the four areas that underpin my theoretical framework (Content, Time, Joint Intellectual Action, and Impact).

5.2 Sociocultural Analysis of Feedback Discourse

One element of my ASCDA approach involved the generation and application of codes to the feedback corpus to provide a picture of the whole data. Table 5.7 shows that I applied a total of 13231 discourse codes to the data. These data are shown as a split across the Chemistry [C], Economics [E], and Geography [G] subjects.

Table 5.7: Discourse Themes/Codes across Subjects

SCDA Theme	**Superordinate Theme/Code**	**Subordinate Code[16]**	**C**	**E**	**G**	**Total**
Content	Focus/Interpersonal	(FI) Opening	268	115	440	823
		(FI) Closing	280	144	546	970
		(FI) Disagreement	216	260	395	871
		(FI) Agreement	272	102	350	724
		(FI) Confronting gaps	14	24	49	87
		(FI) Distancing	357	85	347	789
		(FI) Authority	110	185	121	416
		(FI) Social common ground	28	7	133	168
		(FI) Accentuation	56	84	544	684
Content	Giving information	(GI) Directive information	103	78	354	535

[16] FI (Focus/Interpersonal); GI (Giving/Requesting Information); Br (Bridging); Act (Action).

		(GI) Mark principle information	122	157	283	562
		(GI) Privileged information	11	16	14	41
		(GI) General information	284	60	349	693
		(GI) Technical information	80	70	323	473
		(GI) Mark statement	90	33	219	342
		(GI) Location of credit	569	232	602	1403
		(GI) Rationale for credit	241	294	415	950
		(GI) Standards	27	36	54	117
		(GI) Examiner rationale	34	105	122	261
		(GI) Request information (Closed)	42	119	145	306
		(GI) Request information (Open)	13	13	16	42
Joint Intellectual Action	Bridging information	(Br) Internal reference	19	3	58	80
		(Br) External reference	106	161	196	463
		(Br) Historic reference	46	102	293	441
		(Br) Cultural reference	1	2	16	19
		(Br) Balancing perspectives	35	42	114	191
		(Br) Offer support	22	14	53	89
Impact	Action	(Act) Review	36	43	168	247
		(Act) Resolution	65	141	188	394
		(Act) Reification	9	3	38	50

This coded data set allowed me to explore, at a preliminary level, the character of the feedback data to see if there were any obvious patterns that could relate to the areas of interest in my theoretical framework.

Consistent with the word count analyses reported earlier, I was interested in establishing whether the nature of feedback communication across these areas of interest differed between the different subjects, across time, or between examiners with differing levels of experience or familiarity. This involved an analysis of the codes that I generated and applied to the discourse during my data analysis.

To choose an appropriate method for analysing any differences in the mean count of discourse codes for these different groups I again needed to check for assumptions of normality. A Shapiro-Wilk normality test (Table 5.8) showed that the data for each discourse code was not normally distributed ($p = <.001$), meaning that I needed to use a non-parametric test to compare the discourse code data means.

Table 5.8: Shapiro-Wilk Normality Test for the Discourse Coding Data

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Opening [FI]	.278	991	.000	.782	991	.000
Closing [FI]	.298	991	.000	.784	991	.000
Disagree [FI]	.363	991	.000	.465	991	.000
Agree [FI]	.280	991	.000	.601	991	.000
Confrontation [FI]	.529	991	.000	.266	991	.000
Distance [FI]	.278	991	.000	.728	991	.000
Authority [FI]	.419	991	.000	.522	991	.000
Social Common Ground [FI]	.501	991	.000	.309	991	.000
Accentuation [FI]	.378	991	.000	.492	991	.000
Giving Directive [GI]	.366	991	.000	.587	991	.000
Mark Principle [GI]	.361	991	.000	.552	991	.000
Privileged Information [GI]	.530	991	.000	.147	991	.000
General Information [GI]	.298	991	.000	.670	991	.000
Technical Information [GI]	.398	991	.000	.612	991	.000
State Marks [GI]	.452	991	.000	.448	991	.000
Locate Marks [GI]	.316	991	.000	.502	991	.000
Give Rationale for Marks [GI]	.292	991	.000	.570	991	.000
GIve Standards [GI]	.513	991	.000	.250	991	.000
Examiner Rationalises [GI]	.444	991	.000	.310	991	.000
Closed Request [GI]	.433	991	.000	.464	991	.000
Open Request [GI]	.533	991	.000	.160	991	.000
Internal Bridge [Br}	.524	991	.000	.208	991	.000
External Bridge [Br]	.392	991	.000	.460	991	.000
Historic Bridge [Br]	.389	991	.000	.525	991	.000
Cultural Bridge [Br]	.529	991	.000	.078	991	.000
Balancing Perspectives [Br]	.495	991	.000	.397	991	.000
Offer Support [Br]	.531	991	.000	.304	991	.000
Review [Act]	.468	991	.000	.409	991	.000
Resolution [Act]	.399	991	.000	.350	991	.000
Reification [Act]	.538	991	.000	.213	991	.000

a. Lilliefors Significance Correction

As with my word count analysis, I chose to use the Kruskal Wallis H nonparametric test. This test can be used for comparing the means of more than two samples (e.g. the word count differences across three subjects in my sample or across the four quartiles of feedback time). The data satisfied the other assumptions for the test to be used (i.e. my dependent variable was continuous, there was independence of observations, and the distribution for each group of the independent variable had the same shape).

First I analysed the mean discourse code count according to Subject (Chemistry [C], Economics [E], Geography) [G] (Table 5.9).

Table 5.9: Kruskal Wallis H Nonparametric Test (Subject Discourse Coding)

Ranks

	Subject	N	Mean Rank
Distance [FI]	C	248	687.42
	E	371	329.79
	G	372	534.15
	Total	991	
Locate Marks [GI]	C	248	538.40
	E	371	442.88
	G	372	520.71
	Total	991	
Giving Directive [GI]	C	248	454.44
	E	371	407.80
	G	372	611.67
	Total	991	
Technical Information [GI]	C	248	449.34
	E	371	412.13
	G	372	610.75
	Total	991	

Test Statistics[a,b]

	Distance [FI]	Locate Marks [GI]	Giving Directive [GI]	Technical Information [GI]
Chi-Square	288.117	25.327	144.386	147.297
df	2	2	2	2
Asymp. Sig.	.000	.000	.000	.000

a. Kruskal Wallis Test

b. Grouping Variable: Subject

This analysis showed that there was a significant difference in the discourse codes across the three subjects for four codes: *Distancing* (H = 288.12, p = <.001, η^2 = 0.29); *Locate Marks* (H = 25.33, p = <.001, η^2 = 0.03); *Giving Directives* (H = 144.39, p = <.001, η^2 = 0.15), and *Technical Information* (H = 147.30, p = <.001, η^2 = 0.15). Effect size calculations showed that *Distancing* accounted for the greatest amount of variance (η^2 = 0.29).

To establish the location of this significance finding I carried out Dunn's pairwise post hoc tests for the three pairs of groups for each discourse code.

- There was very strong evidence (p <.001, adjusted using the Bonferroni correction) that Geography and Economics messages used more *Distancing* than Chemistry messages (*Mdn* = 1.0 compared with *Mdn* = 0.0);
- There was very strong evidence (p <.001, adjusted using the Bonferroni correction) that Chemistry messages used more *Locate Marks* than Geography and Economics messages (*Mdn* = 0.5 compared with *Mdn* = 0.0);
- There was very strong evidence (p <.001, adjusted using the Bonferroni correction) that Economics messages used more *Giving Directives* than Geography and Chemistry messages (*Mdn* = 1.0 compared with *Mdn* = 0.0);
- There was very strong evidence (p <.001, adjusted using the Bonferroni correction) that Economics messages used more *Technical Information* than Geography and Chemistry messages (*Mdn* = 1.0 compared with *Mdn* = 0.0);

In the next analysis I wanted to analyse the feedback codes according to examiner experience and familiarity. As these categories of data were dichotomous I chose to use a Mann-Whitney U nonparametric test. This test can be used for comparing the mean rankings of two samples that are not normally distributed (e.g. the discourse code count of feedback messages between TLs and new and experienced examiners, or between TLs and familiar and unfamiliar examiners). The data satisfied the other assumptions for the test to be used (i.e. my dependent variable was continuous, there was independence of observations, and the distribution for each group of the independent variable had the same shape).

Table 5.10: Mann-Whitney U Nonparametric Test (Experienced/New Examiner Discourse Coding)

Ranks

	Experienced Examiner	N	Mean Rank	Sum of Ranks
Distance [FI]	Yes	524	459.93	241002.50
	No	467	536.47	250533.50
	Total	991		

Test Statistics[a]

	Distance [FI]
Mann-Whitney U	103452.500
Wilcoxon W	241002.500
Z	-4.579
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Experienced Examiner

My analysis (Table 5.10) indicated that *Distancing* was more frequently represented in the feedback communication between TLs and new examiners (U = 103452.500, p = <.001, r = 0.15), with the median count for experienced examiners being 0.0 but for new examiners being 1.0 codes. My analysis (Table 5.11) also indicated that *Distancing* was more frequently represented in the feedback communication between TLs and unfamiliar examiners (U = 110439.500, p = .008, r = 0.08) with the median count for familiar examiners being 0.0 but for unfamiliar examiners being 1.0 codes.

Table 5.11: Mann-Whitney U Nonparametric Test (Unfamiliar/Familiar Examiner Discourse Coding)

Ranks

	Prior Professional Relationship	N	Mean Rank	Sum of Ranks
Distance [FI]	Yes	442	471.36	208342.50
	No	549	515.84	283193.50
	Total	991		

Test Statistics[a]

	Distance [FI]
Mann-Whitney U	110439.500
Wilcoxon W	208342.500
Z	-2.649
Asymp. Sig. (2-tailed)	.008

a. Grouping Variable: Prior Professional Relationship

The respective effect size calculations for these discourse codes showed that they each accounted for a small amount of between group variance, with the effect being greater for new than for unfamiliar examiners.

I discuss each of these findings in the next section, where I describe the findings from my qualitative analysis of the feedback according to each of my areas of interest (*Content, Time, Joint Intellectual Action*, and *Impact*). To focus my analysis I use the outcomes of my coding analysis (outlined above) to guide my exploration of the feedback data. Given the space limitations of this thesis I focus my discussion on the clusters of codes that accounted for greater than 75% of the discourse coding in each of my areas of interest.

5.2.1 Content

According to the theoretical framework on which my coding schedule was based, content has both interactional and transactional elements, which are heavily interconnected. These elements are represented broadly in my *Interpersonal Focus* and my *Giving Information* themes.

In this section I initially deal with the Interpersonal Focus content theme, where six codes (being applied 4861 times in total) accounted for 78% of my coding in this area. These codes were *Opening*, *Closing*, *Disagreement*, *Agreement*, *Distancing*, and *Accentuation*. At the same time, I acknowledge that dealing with discourse in this way is to some extent overly simplistic; since the interactional elements of communication are often interleaved with transactional content (i.e. studying the style of communication assumes that *something* is being communicated). As a consequence, my following analyses take the interactional element of discourse as the primary focus but also use transactional content where it is required to contextualise the analysis.

5.2.1.1 Opening and Closing

Nearly all of the feedback messages included an opening and a closing section Analyses suggest that the participants used openings and closings in a purpose driven way so as to achieve particular effects. This interpretation is perhaps most clearly demonstrated in the case of TL2 who generally did not use personalised greetings in feedback message opening/closing. This approach is revealed in the two feedback extracts below (Figure 5.1). In each of these the TL includes an orienting title (e.g. 'Standardisation Script 3') followed by information about marking guidance.

```
Standardisation Script 3

1b - choice/decision mark can only be awarded if given up/sacrifice
mark has been awarded.
3 - tick should be on high quality.
4b - mark should be awarded for correctly identifying that the price
has decreased.

[E9 25.5 1657]
```

```
Standardisation 1 Script 5
4c - we only gave one mark for the first reason as more substitutes
would shift the demand curve inwards, not outwards.

6b - note the diagram doesn't get L3 - the explanation of the
diagram does. Mark spot on.

[E10 27.5 1853]
```

Figure 5.1: Non-personalised Feedback Extracts

The TL rationalised superfluous communication as being a potential distraction from the point of the message:

> *I don't need to waste [words], the potency of the message goes in the more words you use in my opinion.*
>
> *[Y1 TL2 [E8E11E12] SR Interview 75]*

This dominant approach contrasts with another piece of feedback that this TL sent to an examiner. In the extract below (Figure 5.2) the TL personalises the message with an informal greeting (`Hi`) and a closing, which includes the TL's name and thanks the examiner for their attention.

From: TL: 6/9/2014 6:16:15 PM

Hi *Examiner name*,

This is out by quite a bit again but I can see the calls you have made (except for one mistake). Can you please carefully review the below?

3 - disadvantage is not clearly resulting from or linked to specialisation, so no credit given.

4c - no identification mark but two explanation marks for substitute good becoming available.

5aii - if income falls demand barely changes gets a mark.

6a - you have given BOD but typed zero into the box. It got a mark.

6b - because they say a tax causes prices to rise it gets into L2. I know it is followed by a shift in demand but they aren't saying the demand shift causes the price rise so they can be credited.

Thanks,

TL name

[E11 6.9 1816]

Figure 5.2: Personalised Feedback Extract

When asked about this apparent shift in style, the TL explained that they related to the experience of receiving blunt feedback and that this could have a demotivating impact on the recipient. The TL also alludes to the human relationship that helps to temper the communication and reception of bad news:

> *I have been at the other end of this and you really just want to know where you have gone wrong… I couldn't give all this, which he'll know will be out of tolerance, and I honestly don't want to dishearten him because I have been*

sending this message and to be fair the message was, you know, 'You are out'... There is a human interaction there whereas actually when you are giving just very straight forward feedback 'this is right, this is wrong', you don't need as much as that but you need to be a bit softer [here] I guess.

[Y1 TL2 [E8E11E12] SR Interview 69-71]

5.2.1.2 Agreement and Disagreement

Disagreement was very common throughout the discourse. It was not possible to use statistical testing to establish whether it was more common than agreement because the coding process is likely to have under-reported the quantity of disagreement that existed in comparison to agreement. The clustering of information for a number of different items in a message, which is a noted characteristic of email communication (for example, see Giordan, 2003), meant that coding tended to hide individual elements of disagreement. For example, one message that conveyed disagreement around 10 misaligned items was coded as one instance of disagreement. If there was an indication of agreement within that message, this was also coded as one instance of agreement.

Therefore, another measure of disagreement and agreement was calculated which took into consideration the number of paragraphs across the whole corpus where either disagreement or agreement was coded. These analyses showed that disagreement was signalled in 3781 paragraphs whilst agreement was signalled in only 1099 paragraphs. At the same time, this form of analysis did not consider the quantity of negative or positive feedback that prevailed in each of these paragraphs. A final analysis was carried out to consider the percentage of each message that was coded as either disagreement or agreement. Table 5.12 shows the balance of agreement or disagreement information across the feedback messages in the whole data sample. These data are also displayed in Figure 5.3.

Table 5.12: Disagreement and Agreement Content in Feedback Messages

	Messages (n)	**Messages (%)**
Disagreement is greater than agreement *[Complete disagreement (no indication of agreement)]*	693 *[147]*	69.9 *[14.8]*
Agreement is greater than disagreement	178	18.0
No agreement or disagreement indicated	120	12.1
Total	991	100.0

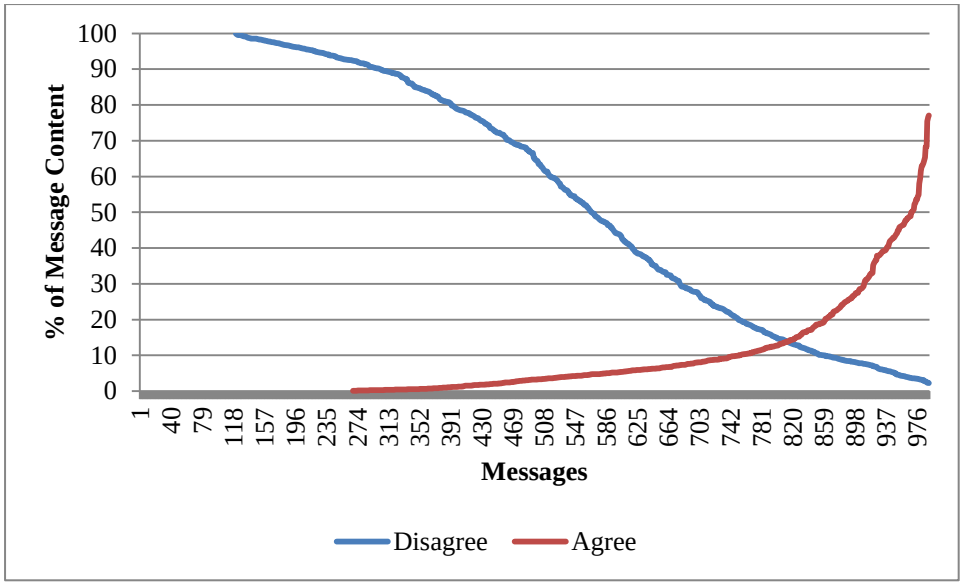

Figure 5.3: Disagreement and Agreement Content in Feedback Messages

This analysis shows that nearly 70% of the messages included more disagreement than agreement (compared to 18% that contained more agreement than disagreement). It is also noteworthy that around 15% of all of the messages contained only disagreement. This finding accords with others that suggest that feedback is mainly a negative discourse (King et al., 2008; Yelland, 2011). An example of feedback that contains only disagreement is shown below (Figure 5.4):

Hi *Examiner name*
Many thanks for your 10 stands. There are a few that are outside of tolerance and will be sent back for checking, please.

The following feedback is as follows:
ID: 648979720: 1d(ii) is fine for the mark
4b(ii) Loses both marks - incorrect species which negates the 2nd mark.
5a I was also penalised for this one - the n on the polymer needs to be a subscript.
6a(i) The higher T curve crosses the lower T curve and does not score.
8b The C2H5O+ fragment scores (at m/z = 45)
ID: 648686202: 1c is 1,1-dibromoethane - zero
1f(i) Is not a NR (No response) = zero
4a The n value is not 0.3 - but 0.03.
5a The monomer doesn't have an n for a balanced equation.
6a(i) The temperatures are not BOD - they must be annotated in some way.

If you could have a look at the five scripts I have returned, remark them and send them back to me, please.

I can then approve your marking.

Best regards
TL name

[E3 21.6 1416]

Figure 5.4: Feedback Containing Disagreement

According to my theoretical framework, the negative content represented in much of the feedback discourse would be expected to have consequences for the way that the participants constructed their interaction. This feedback has disputational features (i.e. disagreement using implied assertion and individualised decision making), which Mercer (2000) and Littleton & Mercer (2013) observe is least productive in terms of affording effective learning (in comparison with Cumulative and Exploratory discourse that encourages shared, collective thinking). My theoretical framework also suggests that negative feedback communication would also present a concern, and that the participants might need to mitigate the effects of disputation in order to maintain on-going professional relations (and afford further learning). From my perspective, this mitigation action would perform a form of articulation work

(Strauss, 1985; Schmidt, 1994, 2011) to coordinate interconnected work across individuals. Morand (2000) notes that criticising, disagreeing and interrupting represent potential face threat, and that the effects of this can be minimised by the use of negative politeness strategies as a form of redressive facework to reduce friction (Culpeper & Haugh, 2014; Lakoff, 1979).

My analyses suggested that the participating examiners were using language to transform potentially disputational discourse to more cumulative discourse. This transformation appeared to include two strategies: (a) *shifting the focus of discourse from strong (potential) disagreement towards weak disagreement*, and (b*) maintaining face* (Goffman, 1955) throughout the interaction. The working hypothesis is that the outcome of the transformation would be to extend the (albeit limited) exploratory potential of between examiner feedback discourse to help the participants to attain aligned perspectives. Such an outcome implicates a type of *Joint Intellectual Action*, as the participants indicate recognition of the 'thought position' of their communicating partner. For me, 'thought position' is similar to the concept of 'perspective', except that it implies a degree of public accountability. This is because participants can use feedback discourse to make their personally held perspective visible to others.

This element links my thinking to the concept of accountable talk (Michaels, O'Connor, & Resnick, 2008). Although my data is not formed exclusively of 'talk', the concept of accountability being important for the development of learning discourse is a common underpinning feature for both. According to Michaels et al. (2008), participants in discourse need to be accountable to others' learning needs (building on each other's ideas), be accountable to the standards of reasoning (the logic of an argument), and to be accountable to the nature of the knowledge that is common to the participants. This final accountability is considered to be the most difficult to achieve because it requires the participants to engage with each other around the content of talk so as to change the substance of the common knowledge upon which discourse is grounded (rather than the style or dynamics of the interaction) (Alexander, 2010).

In summary, I argue that the concept of *Joint Intellectual Action* can be seen as a form of dialogue. The concept of dialogue is contested, but in the context of this study it describes the way that remote participants come to engage with each other's thinking in the pursuit of coordinated actions. Whilst it might be argued that dialogue is intellectual engagement

through language for the purpose of co-constructing new knowledge, here the purpose is better expressed as orientating perspectives to construct common knowledge on which the participants can act.

The sections below use extracts from feedback transcripts to exemplify how the TLs structured their feedback to transform disputational to more cumulative discourse. This transformation involved two broad strategies; (a) shifting the focus of discourse, and (b), face management. The spoken feedback transcriptions include annotation indicators to help the process of inferring the nature of the participants' interactions. These annotations are adapted from Jefferson (2004) and Martin & Rouncefield (2003) and are presented in Chapter 4 and Appendix G.

5.2.1.3.1 Shifting the Focus of Discourse

There is evidence from TL telephone feedback that they were managing disagreement in order to minimise potential examiner disaffection and to ensure that the examiner continued to mark. Feedback is a genre that is often closely aligned with the delivery of bad news (Yelland, 2011), and there is evidence that the TLs were organising the content of their communication to shift the focus of discourse in a number of ways.

5.2.1.3.1.1 Underplaying Differences

In extract E1 619.1200 (Figure 5.5) the TL manages engagement through underplaying the seriousness of the disagreement (line 005) [little bit]. The use of apology also reinforces the dispreference related to giving bad news (lines 003, 004, 008, 009) [sorry] [unfortunately] [afraid]. The TL shifts the focus of the conversation (line 012) emphasising [now] before the examiner can dwell on the disagreement. This is a shift towards positive help following the delivery of negative news.

```
001   TL    Hi Examiner name?
002   EX    Yes speaking
003   TL    TL name, Hi, I'm s:: I'm sorry I had to send you back
004         another set of scripts but erm unfortunately with the
005         first batch being a little bit over the [limit
006   EX                                            [I was a bit
007         yeah]
008   TL    Yah], I'm afraid I've got to send them back so you'll
009         have to do another set I'm afraid? And then submit those
010         before we can [be]
011   EX                  [Yeah]
012   TL    Up and running. Now is there anything in particular,
013         you, you said something

[E1 619.1200]
TL [Team leader] EX [Examiner]
```

Figure 5.5: Underplaying Difference in Discourse

TLs also displayed appreciation and de-personalised the focus of content to mitigate the potential effects of negative messaging. Extract E1 621.1600 (Figure 5.6) starts with the examiner expressing doubts about their own ability (lines 010-011/013/021-024) [mistakes] [disappointing] [miffed] [had such a problem] [disappointed]. The TL manages engagement through underplaying the seriousness of the disagreement (006/015/037) [little bit] [bits and pieces]. The TL also foregrounds the bad news by thanking the examiner (line 003), which sets the scene more positively than starting with bad news. The role of the use of politeness devices to reduce friction has been noted elsewhere (Lakoff, 1979, p. 64). The TL finishes by reinforcing that the problem is complex (line 025) to reassure the examiner by implying that this is a general rather than a personal issue. The TL then shifts the focus of the conversation (line 019) towards the need for the examiner to attend to feedback as a remedy to disagreement. This is a shift towards positive help following the delivery of negative news.

```
003   TL    Hi, have you got my, have you, thank you for your second
004         batch basically
005   EX    Ok
006   TL    Four of them were a little bit outside of tolerance so
007         I've sent them back if you could re-look at my feedback,
008         amend them as soon as possible, put them through, and
009         I'll approve you
010   EX    Ok, you think so? Because there's still a lot of
011         mistakes by the looks of it
012   TL    Erm [yes]
013   EX        [Quite] disappointing really
014   TL    It's but it's similar sorts of things, y'know some of
015         the bits and pieces which are on the mark scheme which
016         you need to highlight with a pen or something so that
017         they stick out
018   EX    Yeah>
019   TL    Erm, but yeah, if you have a look at my feedback, you've
020         got the feedback I presume?
021   EX    Yeah I have I'm just a bit [laugh] miffed by this mark
022         scheme because I've marked for years and I've never had
023         such a problem, I'm just a bit miffed I really am a bit
024         disappointed really
025   TL    Yes it's quite a complex one
[…]
034   EX    I can't I can't do this if that's the case
035   TL    No you can, you can, honestly
036   EX    (laugh)
037   TL    It's just little bits and pieces, it's only little
038         [bits]

[E1 621.1600]
```

Figure 5.6: Giving Thanks and De-personalisation in Discourse

Steering communication around problematic areas is also evidenced in extract E9 25.5 T 83-95 (Figure 5.7). In the opening turn (lines 001-003), the TL plays down their difference with the examiner [some of the time there are judgments, the marks we've given are not necessarily the perfect marks] whilst also holding onto a position of authoritative judgement [what you've given… I'd say 'fair enough']. The importance of the difference between the TL and the examiner is also reduced by the implied suggestion that any discrepancy between judgements is understandable to some extent, given the complexity of the performance being assessed (lines 006-007) [this is a tricky script].

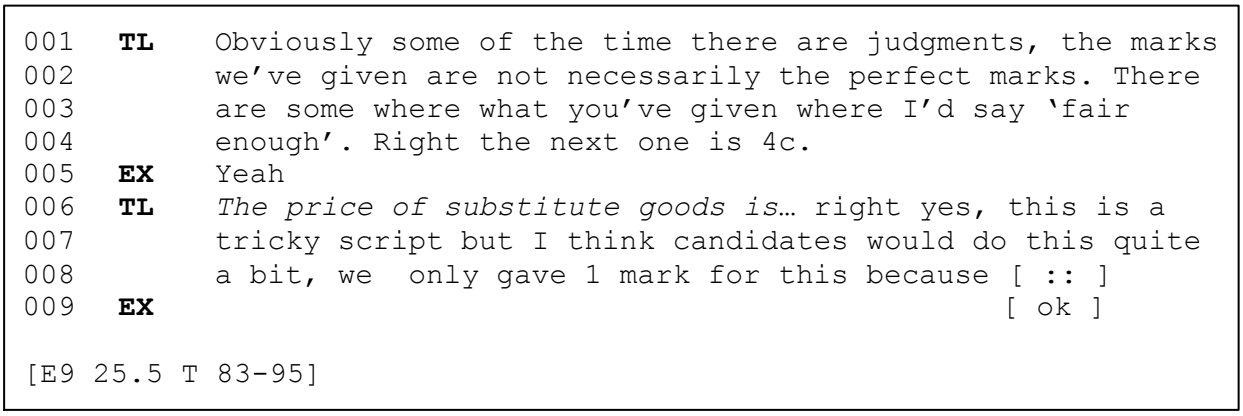

Figure 5.7: Steering Communication around Problematic Areas

5.2.1.3.1.2 Avoiding Problematic Discussion

In extract E2 a 617.0001 (Figure 5.8), an examiner doubts their own competence and responds to the TL's questions with emphasised dispreference markers (line 009) [have] (line 011) [I wouldn't say running]. This suggests that the examiner has problems with the mark scheme. This is reinforced with the emphasis placed on [much] (line 014). The TL response in all of these cases is to steer the discussion towards a positive interaction (line 010) [good], downplaying the complexity of the mark scheme (line 015) [little interesting] (line 020) [you'll cope]. During the interaction, the TL does not take up the opportunity to enquire about what the examiner finds complex, instead offering a view on why Question 7 may not prove to be such a problem (line 017) [But].

```
008   TL    Good. Have you downloaded the mark scheme yet?
009   EX    I have
010   TL    Oh good, fine, so you're up and running
011   EX    Sort of I wouldn't say [running]
012   TL                           [no] I know it's quite a marathon
013         I'm afraid this one this year
014   EX    But erm not much that seems very awkward:: not much
015   TL    No I think 7 can be a little interesting [but]
016   EX                                             [yes]
017   TL    But most candidates by the time they get there are quite
018         tired and so their answers aren't brilliant. There's a
019         few little bits and pieces there to watch out for. I'm
020         sure you'll cope

[E2 a 617.0001]
```

Figure 5.8: Avoiding Problematic Discussion

5.2.1.3.2 Maintaining Face

Analyses of telephone and email feedback suggest that TLs managed face saving through their communication. In these messages the TLs used discourse content to downgrade the importance of their disagreements with examiners, to demonstrate and invite reciprocation, and to demonstrate symmetrical relations so as to undermine the dominance of the hierarchic relations implicit to the participants' roles.

5.2.1.3.2.1 Softening Definitiveness through Modals

In extract E3 18.6 1456 (Figure 5.9), the TL's responses are embedded within the examiner's comments (indicated in red font). This is a cohesion building strategy adopted by the TL that helps to ensure that their responses are seen to address directly the issues raised by the examiner. The examiner acknowledges and changes their approach to marking whilst acknowledging the apparent mixed messages from the TL (lines 006-008). The disagreement is weakly stated, downgrading the disagreement by reducing the definition of the TL's own original marking position (line 014) [`I'm afraid so – possibly`]. The TL also softens the definitive nature of the responses through the use of modals [`may have`] [`if`]

and adverb [possibly] (lines 010/011/014), which reduces the implication that the examiner is completely incorrect.

```
TL: 18.6 1456
001   Hi Examiner name,
002   Thank you for the feedback, I have amended the 2 you sent back
003   to me.
004   2 queries:
005   ID 649581302 - Q1 g ii - MS says 'it' should be assumed to mean
006   cyclohexane. Do they still need to have written cyclohexane
007   somewhere in their answer to get the mark? I accepted 'It burns
008   more effectively'.
009   I cannot find the comment re: 'assumed to be cyclohexane' in
010   the mark scheme - it may have appeared in the practice scripts
011   by the sound of it, and was incorrect if it did.
012   ID 649661411 - Q2b - do they get the mark even though 'curly'
013   arrow is almost straight?
014   I am afraid so - possibly a little generous.
015   I will look through the other 5 and send over.
016   Thanks.
017   TL name

[E3 18.6 1456]
```

Figure 5.9: Softening Definitiveness through Modals

5.2.1.3.2.2 Recognising the Participant's Voice

In the next extract (Figure 5.10), comprising three messages over two hours, the TL recognises the perspective of the examiner (line 001) [I can see why you awarded three] before outlining their own perspective, marked with a contrastive adjective [but]. Other contrastive adjectives (e.g. 'although') were also found in TL discourse. The recognition of examiner perspective reduces the significance of the disagreement whilst also bridging the views of the participants. This significance reduction is completed with another downgrade (line 011) [a matter of opinion!].

```
TL: 602.1922
001   6 - we gave zero. I can see why you awarded three but we felt
002   the idea that heating prices would fall was not well connected
003   enough to the question in hand - ie. they didn't say why that
004   came about and how it was related to what was going on in the
005   agricultural market.

Ex: 1923
006   6 This implies that the ideas in the MS that are being accepted
007   as creditable must form part of an argument and that therefore
008   they cannot be accepted in isolation, unless they are one or
009   more of the basic ideas regarding supply and demand.  Is that
010   right?

TL: 2101
011   Understood, this is definitely one that is a matter of opinion!

[E6_602.1922/1923/2101_Q4]
```

Figure 5.10: Recognising the Participant's Voice

The participants demonstrated and invited reciprocation in their communication, reinforcing the interconnected nature of their work. In extract E12_707.2111/708.1149/708.1149_Q2 (Figure 5.11), which was carried out over two days, the use of thanks, initiated by the examiner, introduces a politeness to the communication that is reciprocated in the response from the TL. This mirrors the patterns noticed in ordinary talk, where the refusal to reciprocate politeness markers is considered to be a dispreference as it flouts the conversational contract (e.g. Fraser & Nolen, 2009).

```
TL: 707.2111
001   Hi Examiner name; 4174; 1021 6771…
002   Finish the others tomorrow I will have a look and then I should
003   be able to completely authorise you and you can then work
004   unhindered by me!
005   Best wishes
006   TL name

Ex: 708.1149
007   Hi TL name,
008   Thanks once again for feeding back to me on the latest batch of
009   scripts. I'll do a few more today and let you know when they're
010   ready.
011   Best wishes,
012   Examiner name

Ex: 708.1149
013   Thanks TL name

[E12_707.2111/708.1149/708.1149_Q2]
```

Figure 5.11: Politeness and Reciprocation

5.2.1.3.2.3 Using Apology to Reduce Face Threat

Similarly, extract E9_602.1955/603.1924/604.1134_Q4 (Figure 5.12) shows how politeness markers are used to open and close interactions (lines 001/002) [sorry] [thanks]. Apology is used by the examiner to preserve the negative face of the TL (reducing intrusion into their professional space). In the extract the examiner also uses the final turn to reassure the TL that their information had been acted upon. This is a reciprocal action deriving from the interaction and serves to validate the TL's communication.

```
Ex: 602.1955
001   Sorry TL name another question, Q8 p10 - is this
002   enough for L3B2? - thanks Examiner name.

TL: 603.1924
003   "Decreasing welfare foregone in the market" is TV - there would
004   really need to be an explanation of WHY (ie. underconsumption
005   eliminated) or a more technical reference to increased welfare
006   (ie. improves allocative efficiency).

Ex: 604.1134
007   OK - the candidate referenced it later on so gave it 14...
008   thanks, Examiner name.

[E9_602.1955/603.1924/604.1134_Q4]
```

Figure 5.12: Using Apology to Reduce Face Threat

I have used the concept of 'articulation', borrowed from Strauss (1985) and adapted by Schmidt (1994, 2011), to simultaneously reference the notion of 'expressing' and the act of 'coordinating interconnected work across individuals', and argue that feedback can be used to perform both functions (Johnson, 2015). It has been noted elsewhere that feedback has both technical and social dimensions (Adcroft, 2011), and one such social dimension appears to be the use of feedback to manage relations. This use coheres with the observation that an important macro-function of language is the effective management of relationships (Brown & Yule, 1983).

A cue to how feedback is used to socially manage relations is found around the use of politeness markers or hedges. Keyword analyses show how the act of giving directives or sharing a personal perspective linked with the use of politeness markers and hedges. These appear to break down the potential for the discourse to be construed by the receiver as being overly negative, thus encouraging discourse to continue. The example below (Figure 5.13) shows how in the 1st/2nd Quartile the target word [not] collocates with the keyword [please] so that TLs can give directives without appearing to be confrontational. The target word also collocates with hedges, such as [really] and [quite], where a TL is communicating their perspective to an examiner. The effect of this collocation is to downgrade the level of definition of the perspective being shared. The effect of this is to convey something of the complexity of the decision-making process to the examiner (e.g. that

assessment decisions include a balance of evidence or a degree of subjective appraisal). An implication of this use of language is that it can reduce the degree of asymmetry in the discourse, so that any disagreement between the participants is considered to be marginal.

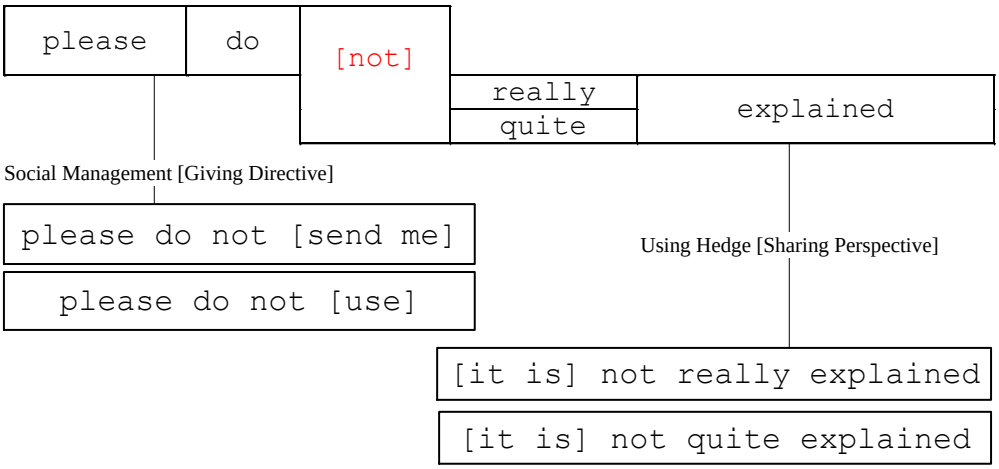

Figure 5.13: Directives and Confrontation

The use of apology is strategy for achieving distancing. A closer look at the use of apology also gives insights into how the participants maintained order and on-going professional interaction in a context where Face Threatening Acts (FTAs) are present. A search of the whole corpus using the search terms ([sorry][apol*]) located 142 instances of apologetic utterances, with a relatively even split between the number of TL apologies (n=70) and examiner apologies (n=72). Additional analysis showed that apology was overwhelmingly used to support the hearer's negative face (125 apologies were coded as supporting the hearer's negative face compared with only five that were used to support the hearer's positive face).

Taken together, these analyses suggest that the achievement of resolution relies on the participants maintaining on-going professional discourse. In turn, this discourse involves the participants using social distancing strategies, such as politeness and apology, to mitigate the potentially negative effects that disagreement may have on the continued discourse.

In this way, apology and politeness give clues to the relational fabric that is established and maintained by the participants. Examples from the feedback corpus (Figure 5.14) show that the participants prefer to create a respectful distance with each other. This distancing involves a dispreference for intrusion on discourse partners. It also appears that the hierarchic nature of the professional environment influences the interaction. In the case of the TLs, it appears

that there is a desire to support the professional status of examiners, with intrusion on the examiner being potentially interpreted as a form of professional undermining.

```
Sorry this feedback is a bit lengthy
[E18 6 6 1122]

I am sorry that this is so long and please be patient and digest it?
[E14 14 6 2015]

Sorry - but this year it appears to be a somewhat protracted
process.
[E3 18 6 1525]
```

Figure 5.14: Distancing in Feedback

This dynamic is also reflected in the tendency for examiners to self-deprecate when they reflect on their mistakes (Figure 5.15).

```
Sorry about this - some careless mistakes here.
[E5 26 6 0729]

Thanks for the feedback and apologies for my errors.
[E3 19 6 2125]

Sorry! I'll go back over the script to see where I went wrong
Thanks!
[E15 10 6 1314]

Sorry to be a pain.
[E17 29 5 1835]

Please accept my sincerest apologies. You must think me awfully dim,
but I think that the penny has finally dropped.
[E24_524 0827_Q2]
```

Figure 5.15: Self-deprecation in Feedback

Examiner intrusion on the TL appears to take the form of requests for additional information. Again, examiners tend to convey dispreference through apology when asking TLs to respond to their interaction (Figure 5.16).

```
Thanks and apologies for another question.
[E29_531.1448_Q2]

Sorry to disturb you again.
[E12 1.6 1602]

Sorry for asking again Ben but is the second L3B2 point enough.
[E27_531.1142_Q3]
```

Figure 5.16: Dispreference and Apology in Feedback

In addition to the maintenance of the hierarchic relations that structure the interactions, the time critical nature of the assessment task is also implicit in the discourse. The prevalence of apology suggests that the participants recognise the disruptive cost of unnecessary intrusion on tight work schedules (Figure 5.17).

```
Sorry for the additional work I have caused you and thank you again
for your help and support.
[E17 7.6 2227]
```

Figure 5.17: Intrusion Reference in Feedback

5.2.1.3.2.4 Highlighting Reciprocity

At times the TL used feedback to demonstrate the symmetrical nature of their relations with the examiner, therefore undermining the formal, established hierarchic relations. Extract E10_531.1009/531.1040/531.1043_Q1 (Figure 5.18) shows a shift in purpose of the feedback, from castigation (001) to support (lines 008-009).

```
TL: 531.1009
001   […] Q3c – Examiner name you have not paid attention to my
002   instruction regarding the placement of the L2 symbol on these 9
003   mark questions. It must go on the left hand side against the
004   place where you think that that level has been reached in the
005   writing!
006   You are also still wasting time putting the levels at the end
007   of the essay.  Does this mean that you did not read any of my
008   last email?  I am wondering about that and if I am wasting my
009   time here? […]

Ex: 531.1040
010   Sorry about my aberration regarding the level marking, it will
011   not happen again.
012   Thanks for the very useful comments about highlighting etc.
013   Examiner name

TL: 531.1043
014   Great Examiner name, really pleased at your progress […] They
015   are being very picky about annotations this year as they are
016   important are when scripts sent in for remark and back to
017   centres etc.
018 TL name

[E10_531.1009/531.1040/531.1043_Q1]
```

Figure 5.18: Highlighting Reciprocity

The TL's use of 3rd person [they] to refer to the exam board (line 014) distances the TL from the organisation (and thereby moves them psychologically closer to the examiner). This contrasts with the way that the initial imperative for action (line 001) is personalised (1st and 2nd person) [you/my]. This shift appears to be instigated by the examiner apologising for their actions (line 010).

In the telephone feedback extracts E1 618.0001/ E1 621.1600 (Figure 5.19) the TL's use of 3rd person (lines 012/014) [they] and pronoun [Cambridge] (line 068) also distances the TL from the organisation. This is reinforced by the use of the 2nd person to refer to the TLs (line 007) [we]. Use of the terms [told] (line 009) and [on my back] (068) reinforce the position that the TLs are subject to the same hierarchic pressures as the examiner. The construction of a shared focus on the apparent inconvenience of the exam board's practice is accentuated with the phrasing of [again] (line 015). The symmetry of the participants'

status is reinforced (072) by the use of the nominative plural [we're] to associate the TL and the examiner.

```
007   TL    Indeed yeas, well we've just finished and we're doing our
008         own standardisation scripts at the moment, and getting
009         those sorted out. We were told that OCR were going to
010         email you the mark scheme
011   EX    That's right?
012   TL    And they haven't done so yet
013   EX    That's right? No>
014   TL    They haven't, ok, so if they haven't' done so by the end
015         of today could you email me again and I'll send a copy, my
016         own copy

[E1 618.0001]
```

```
067   TL    But I'd prefer, because otherwise there'll be problems
068         along the line and Cambridge will be on my back [laugh]
069   EX                                                    [I know]
070         I understand all that it's just if I'm going to keep
071         making these mistakes
072   TL    No, well I think we're all going to keep making mistakes I
073         promise you [laugh] I promise you we're all going to keep
074         making mistakes because it's such a very complex mark
075         scheme
076   EX    Yeah , ok TL name

[E1 621.1600]
```

Figure 5.19: Shifting the TL towards the Examiner

Extract E12_09.1207/709.1230/712.2130_Q3 (Figure 5.20) demonstrates how a TL reinforces symmetrical relations in a written response to an examiner-initiated communication. The TL's intention appears to be to seek to hand over the responsibility for marking to the examiner (line 010) [have to let you go]/(line 025) [off you go!]. In so doing, the TL plays down their expertise (lines 014-015) by stating [I am not perfect] and that joint work may reach a better conclusion. This reciprocation serves to empower the examiner, reinforcing positive face, and raising the status of the examiner's judgement to that of the TL.

```
Ex: 09.1207
001   Hi TL name,
002   Your feedback for the latest 4 scripts has been taken on
003   board. I'll continue now with the others. Also, in the future,
004   if I'm really   unsure with a mark I might hold the scripts
005   back and send you a quick email  with the script number(s) and
006   questions if that's ok?
007   Best wishes
008   Examiner name

TL: 709.1230
009   Hi Examiner name
010   Yes, I am going to have to let you go on as I cannot keep
011   checking every script and you are getting more accurate
012   although I am looking for even better?  However, I will keep
013   an eye on you and   contact you if anything is worrying.
014   Please do ask if you are unsure that would be great. I am not
015   perfect of course, far from it, but two heads are often better
016   than one!
…
021   You are going to have to make your own decisions now [name] or
022   we  will be here forever.  Of course the odd question is
023   welcome but the rest you will have to judge now with the help
024   of your training and the mark scheme; you are an approved
025   examiner so off you go!
026   Good luck
027   TL name

[E12_09.1207/709.1230/712.2130_Q3]
```

Figure 5.20: Reinforcing Symmetrical Relations

5.2.1.3.2.5 Building Social Common Ground

At times, the TLs and examiners used feedback messages as an opportunity to share personal information with each other. Extract E6 21.6 T 15-26 (Figure 5.21) demonstrates how an examiner ignores what might be a closure cue (line 016) [`lovely, lovely`] and takes the opportunity to keep the conversation open after a small gap. The TL then returns to the topic of the nice weather (line 020), which the examiner had introduced earlier (line 007). This topic appears to be an area that links the participants, being an opportunity for the participants to share their perspective on work and leisure. This episode is cued by the TL's use of the phrase [`I'm afraid so`] (line 018), which both participants find humorous, and which may imply that work gets in the way of things that would be more personally satisfying.

001 **EX** Hello?
002 **TL** Hello, good morning, that must be *Examiner name*
003 **EX** Hello, speaking
004 **TL** *TL name* [:]
005 **EX** ?Hi there, how are you?
006 **TL** I'm fine, how are you? [:]
007 **EX** Wonderful, beautiful day
008 **TL** (laugh) And you're having time to enjoy it [laugh]
009 **EX** [laugh] I know
010 **TL** Er, everything's fine it's all sent back, you're up and
011 approved, er, so if there's anything on the feedback that
012 you're not happy with just do come back to me
013 **EX** Ok, that's lovely
014 **TL** You're fine, you're up and running, superb
015 **EX** Excellent
016 **TL** Lovely lovely [.]
017 **EX** On with the job then
018 **TL** I'm afraid so [laugh]
019 **EX** [laugh]
020 **TL** but do take time to enjoy the sun, please
021 **EX** Oh yes
022 **TL** We won't have much of it
023 **EX** yeah, well, I mean fortunately I'm retired so I'm not
024 pushed to sort of, it's not as if I'm still teaching and
025 having to do all this in the evenings and weekends
026 **TL** Yeah, true so you might have time to enjoy the sunshine
027 whilst we have it
028 **EX** Exactly, yes, well when I take a break I just go out on
029 the deck and have a cup of tea
029 **TL** (Laugh) very nice, you're making me very envious (laugh)
030 ok, thanks very much, take care, bye now
031 **EX** Ok, bye now

[E6 21.6 T 15-26]

Figure 5.21: Building Social Common Ground

5.2.1.3.2.6 Reinforcing Historical/Interpersonal Relations

In extract E10 TM a 530.0001 (Figure 5.22) the TL appeals to the shared history that the participants have developed and which they find to be personally valuable (lines 031/033) [I'm so glad] [it's just nice]. This appeal follows relatively serious castigation of the examiner, which is acknowledged by the examiner as an acceptance of responsibility (lines 029-030) [no problem at all]. Following the examiner's acceptance of the TL's point, the TL focuses on the interpersonal element of their professional relations, reinforcing

the esteem in which they hold the examiner and their on-going working relationship. Hierarchic roles appear to be clearly implicit to this relationship, as the TL uses an impersonal reference [one] to refer to the examiner (line 033). This shift to formality seems to be in contrast to the use of the 2nd person reference [you're] (line 031), which the examiner then uses as an invitation to respond in an informal humorous manner [you smooth talker you] (line 032).

027	**TL**	But er if you just bear with me now because I can't I
028		can't monitor all of these at the same time [laugh]
029	**EX**	[No problem
030		at all, no problem at all]
031	**TL**	I'm so glad you're with me this year *Examiner name*
032	**EX**	You smooth talker you
033	**TL**	No it's just nice to have one that I've been with before
034		if you see what I mean rather than all the strangers
035		[laugh]
036	**EX**	Yeah
037	**TL**	Ok then
038	**EX**	You want me to go through these three again, you want me
038		to
039	**TL**	Yeah, if you could just finish those thr::ee? and I'll
040		have a look
[E10 TM a 530.0001]		

Figure 5.22: Reinforcing Historical/Interpersonal Relations

5.2.1.3.2.7 Using Metaphors of Non-virtual Relations

In extract E10_624.0952/624.1627/624.1651/624.1700_Q3 (Figure 5.23), captured over four messages, the TL appears to be trying to understand why the examiner has not completed their marking. In this exchange the TL draws on references to physical (i.e. non-virtual) relations through the use of terms such as (line 005) [talk] (lines 014-015) [wandering] [meet]. This accentuates the presence of the initiator as an active social agent in the communication relationship. The TL also appears to indicate something about how they perceive the professional relationship in this extract. References to care for the examiner (line 002) [all is well] (line 005) [worried] (line 012) [thank goodness] emphasises the pastoral dimension of the TL role.

```
TL: 624.0952
001   Hello Examiner name
002   Just enquiring if all is well as you do not seem to be
003   progressing and I hope that nothing is amiss? Have tried to
004   telephone you but just spoke to your answering machine.
005   Worried and hoping that you are OK and will talk to me soon?
006   TL name

Ex: 624.1627
007   I have had my only granddaughter staying for the last few days
008   and as it was her birthday and my wife's I got a bit
009   sidetracked!!
010   Back in the saddle today. Aiming to catch up!!
011   Examiner name

TL: 624.1651
012   Oh thank goodness, how lovely, I hope that you all had a
013   lovely few days. I was getting worried and am so pleased that
014   there was a happy reason! I kept wandering into your area and
015   no one was there to meet me!
016   Thanks for letting me know
017   TL name

Ex: 624.1700
018   Glad to put your mind at rest!!!

[E10_624.0952/624.1627/624.1651/624.1700_Q3]
```

Figure 5.23: Using Metaphors of Non-virtual Relations

5.2.1.3 Distancing

This code links heavily with the preceding section. An important function of distancing is the manipulation of discourse to reinforce participants' face maintenance. It is likely that the ability to judiciously manipulate discourse so that participants have an appropriate degree of proximity to each other also overlaps, to some extent, with the communication of social information. This point is reinforced by Eraut (2000) who argues that 'in order to respond, one has to assume some knowledge of the person one is talking with' (Eraut, 2000, p. 121).

Distancing involves the use of positive and negative politeness. According to the definition by Brown & Levinson (1987), positive politeness reduces the threat to the recipient's positive face by accentuating empathy and common ground between the participants, whilst negative politeness avoids imposition on the recipient's negative face (i.e. the desire to act unimpeded)

by creating respectful distance. In some analyses of TL word use (reported and discussed later in Chapter 6.2.1.2) I show that TL feedback tends to employ negative politeness strategies, and that this differs from observations is some other professional workplaces (Holmes, 2001; Holmes & Stubbe, 2003).

My analyses show that there was very strong evidence that Geography and Economics feedback messages used more distancing strategies, and that these were also most prevalent in communications with new and unfamiliar examiners (Table 5.9; Table 5.10; **Error! Reference source not found.**). These findings cohere to some extent with the expectations of my theoretical framework which suggested that mark-related trouble sources would be most prevalent in subjects where there are more subjectively marked items, and between examiners who have limited prior common ground.

5.2.1.4 Accentuation

Accentuation strategies give an indication of the things that the TLs considered to be important in feedback discourse (and perhaps the things that examiners were not attending sufficiently to). Initial descriptive analyses based on the MAXqda qualitative data analysis outcomes (Table 5.13) suggest that Accentuation was commonly collocated with disagreement. An example of this collocation is shown in Figure 5.24.

```
7c(i) NOT: 'different arrangement of atoms' for 'structural
isomerism'

[E1 19 6 0944]

Q1 all of your marks are one mark less

[e10_626.2102_Q3]
```

Figure 5.24: Accentuation and Disagreement in Feedback

This tentatively suggests that the eradication of disagreement is a primary focus for feedback communication.

Table 5.13: Descriptive Analysis of Accentuation Code Collocations

Code	Collocation with *Accentuation (FI)* code (n)
(FI) Opening	1
(FI) Closing	2
(FI) Disagreement	463
(FI) Agreement	92
(FI) Confronting gaps	11
(FI) Distancing	71
(FI) Authority	35
(FI) Social common ground	19
(GI) Directive information	71
(GI) Mark principle information	98
(GI) Privileged information	6
(GI) General information	108
(GI) Technical information	75
(GI) Mark statement	54
(GI) Location of credit	117
(GI) Rationale for credit	142
(GI) Standards	18
(GI) Examiner rationale	17
(GI) Request information (closed)	22
(GI) Request information (open)	3
(Br) Internal reference	11
(Br) External reference	67
(Br) Historic reference	56
(Br) Cultural reference	3
(Br) Balancing perspectives	15
(Br) Offer support	5
(Act) Review	54
(Act) Resolution	37
(Act) Reification	12

5.2.1.5 Giving/Requesting Information

In this section I deal with the Giving Information content theme, where six codes (being applied 4396 times in total) accounted for 77% of my coding in this area. These codes were *Directive Information, General Information, Technical Information, Mark Statement, Location of Credit,* and *Rational for Credit*.

It is noticeable in the qualitative analyses of the feedback communication that there were individual variances across the subject areas with regards to the ways that feedback information was conveyed. For example, when analysing the characteristics of their feedback

messages it became evident that TL1 and TL3 differed in the extent to which they engaged in Mark Statement or in explaining a Rationale for Credit. This difference is demonstrated in the feedback extracts below that show three interactions from each TL to different examiners.

Morning *examiner name* Many thanks for the latest seed -it was excellent.

1e(i) was correct for the mark. The rest was perfect. (GI) State Mark

Best regards *TL1 name*

[E1 29.6 1036]

Morning *examiner name*

Many thanks for the latest seed which was again excellent.

1e(i) Is a good answer for the mark - some of the answer has encroached on the next question. (GI) State Mark

3e Is the 1st mark and (GI) State Mark

4b(ii) does not attract any marks as the species are incorrect.

The rest was perfect.

Many thanks.

Best regards *TL2 name*

[E2 13.7 1031]

Morning *examiner name*

Thanks for the 1st five scripts - as per my previous message, I have returned the other 5 to be submitted after this feedback.

The marking was fine - there were a couple just outside of tolerance which I will return to you as soon as I have completed the feedback.

ID: 648840477: 2b The 2nd mark can score here. (GI) State Mark

2d We gave the BOD here. (GI) State Mark

8b The Molecular fromula and first fragment both score here. (GI) State Mark

Hope this is helpful - if you could amend the two I have returned, please, and send them back to me.

Then, please, look over the other five before you submit them, and then I will send you the 10 standardisation scripts, after I have given feedback on the second five scripts.

Best regards *TL1 name*

[E3 18.6 1129]

Figure 5.25: Stating Marks in Feedback

In the first three extracts (Figure 5.25) the Chemistry TL primarily communicated the acceptable mark for a performance without any explanation (GI State Mark). This contrasts with the next three extracts from the Geography TL (Figure 5.26) which are noticeable for their use of TL rationalisation (GI Rationale).

Hello *examiner name*

Sorry for the delay getting back to you but I firstly had to wait for my own green light from the PE.

Your first two scripts were quite close to the mark scheme but nevertheless I have detailed them below for you to look at. It is imperative that you take some time to read these notes.

Let us see what is happening. Question 9446

2b, the weathering processes are not well detailed, eg how are the crystals formed etc so only 3 marks, you are correct in level 1 as link to landforms is not shown just 'rock' but only 3 marks not 4. (GI) Rationale
Q2ai only achieved 2 marks because the second set of grid references were wrong and, unlike the first two activities had no other location information so please watch this. (GI) Rationale
3aii You cannot credit 'population' as it was not mentioned in 3ai, please see the notes on double penalty, ie you can credit something mentioned in 3ai here even though it was not correct in 3ai, ie a human factor, but you cannot credit a factor brought new into 3aii without any mention in 3ai. (GI) Rationale
3b PE is not accepting long roots in cold environments because of the frozen ground, accept large or spread out roots though. 2 marks only here. (GI) Rationale
3c This answer only achieved L1 because there is no explanation of how meltwater shaped the landforms/corrie except for where the 'due to the meltwater the glacier can freely move allowing erosion etc. so 3 marks only given. (GI) Rationale

Essay covered in previous paper.

[E14 24.5 1744]

Feedback for last batch, two in tolerance and the other very close, wonderful!

Hi *examiner name* I had a lull in my marking so thought of you, thanks for sending these.

Script 9524

Q2aii This has no annotation on the page. Was given 4 marks

2b Not sure why you stated freeze thaw and carbonation as irrl? Both are weathering processes and, more importantly, both in the mark scheme? So, two points, neither have process detail AND influence of landform so 4 marks, ie not enough for L2. (GI) Rationale

Essay Q7 – AO2 General stated gains, nothing on short-termism so 0 marks here. Plenty of English errors so bottom of level 2 in AO3 so 4 not 5.Script (GI) Rationale

7360

ESSAY Q5 – only one mark out ie 8 given for AO1. Does talk about the role of water which is an indicator for levels here, ie explains diseases and houses swept away. So 8 marks rather than 7. (GI) Rationale

Well what a good set of scripts [examiner name], I am very happy, it has saved me lots of time, Well Done!

[E15 27.5 2353]

Rest of the three standardisation scripts reviewed

Script 1476

1ai and 1aii correct

1b Rainsplash – not undercutting as a result as stated but some linkage given to river cliff landslides, if unclear. Max load not clearly linked to braiding so given 4 marks not 6. (GI) Rationale

1c You were spot on with the level and in fact it was pushed to 9 marks not 8.

3ai and 3aii were correct

3b Two ways well explained so full marks given ie 6 marks instead of 4.

3c Clear role of meltwater in formation of hollow and explained well, all criteria of level 3 met and given 8 marks, not 5. Ie meltwater as an agent of freeze thaw and erosion if explained OK, must explain what meltwater is doing though. (GI) Rationale

Essay Q6

AO1 Sections on actual management used are irrelevant. Section on pollution not relevant to the answer. Ignore them and focus on good areas of the answer – sections on ecological needs, ie little tern are OK. Other need is economic development but actual explanation of this is vague so low L3 given ie 11 marks and you gave 10 so very close. (GI) Rationale

AO2 Limited comments on varied need or at best implied plus only one comment in the conclusion so L1, 2 marks, you gave 3.

AO3 Vague conclusion but all other criteria met so low L3, ie 6 marks, you only gave 4. (GI) Rationale

You are on the right road [examiner name] please look over the other standardisation questions in light of this review before sending them to me.

Thanks *TL3 name*

[E16 26.5 1607]

Figure 5.26: Rationalising in Feedback

These observations cohere with the expectations from my theoretical framework that the nature of the items on the examination paper being marked would be a source of trouble for examiners to resolve through feedback. Other statistical evidence from the code analysis showed that there was very strong evidence (Table 5.9) that Chemistry messages used more restricted forms of discourse (i.e. the *Locate Marks* strategy) compared with the Geography and Economics feedback messages that used more elaborated feedback (e.g. rationalising). This difference, it is suggested, relates to the greater prevalence of objectively marked items on the Chemistry paper compared with the other papers.

My analyses also show how TLs used language to achieve a number of purposes, which contributed to the complexity of feedback communication giving. The examples below (Figure 5.27) show how the keyword [need] is used by a TL to both give a directive and to give a rationale for why marks may be awarded, and the keyword [awarded] is used by TLs to give rationales for their thinking and to indicate where credit may be found in a performance script.

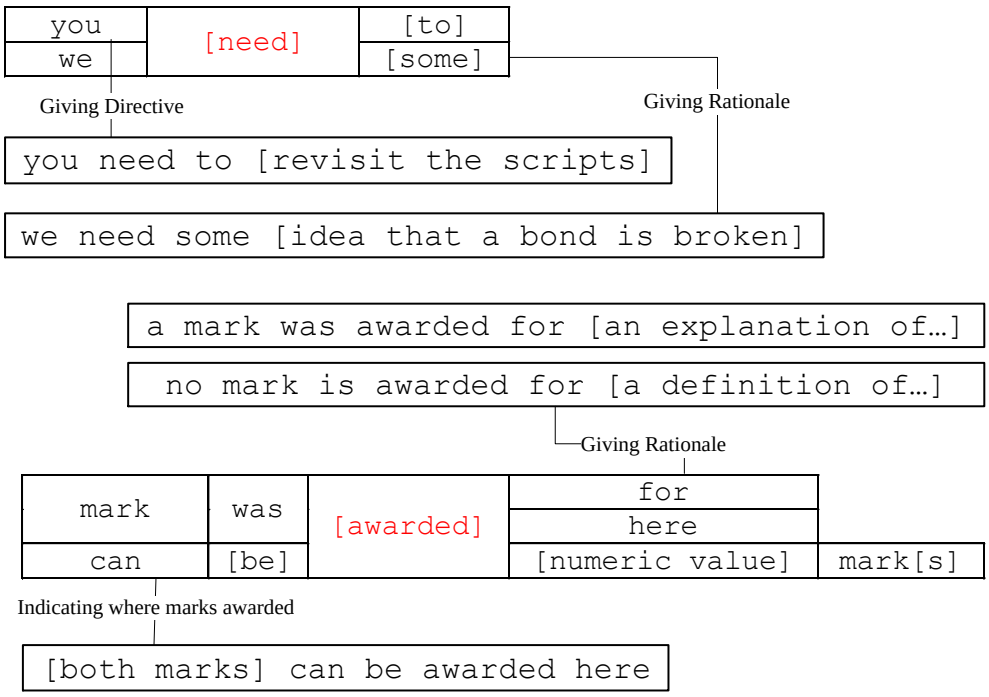

Figure 5.27: Giving Directives and Rationales

As is demonstrated above, *Giving Information* also included the use of language to give directives so as to indicate to a participant what was next required from them. My analyses (outlined earlier) also showed that *Giving Directives* was more common in Economics messages compared with the other subject areas (Table 5.9). My theoretical framework does not adequately help to explain this finding so it requires additional consideration.

The fine detailed analyses that I carried out using ASCDA methodology also uncovered some nuances of communication, including the ways that TLs used language to evaluate an examiner's marking performance. In the example below (Figure 5.28) the 4th Quartile keyword [have] is used by TLs to report their evaluation of the examiner's marking performance. Again, my theoretical framework would suggest that this form of

communication implicates a potential face threat, with emotional consequences, as it questions the professional authority of the examiner to make an independent judgement.

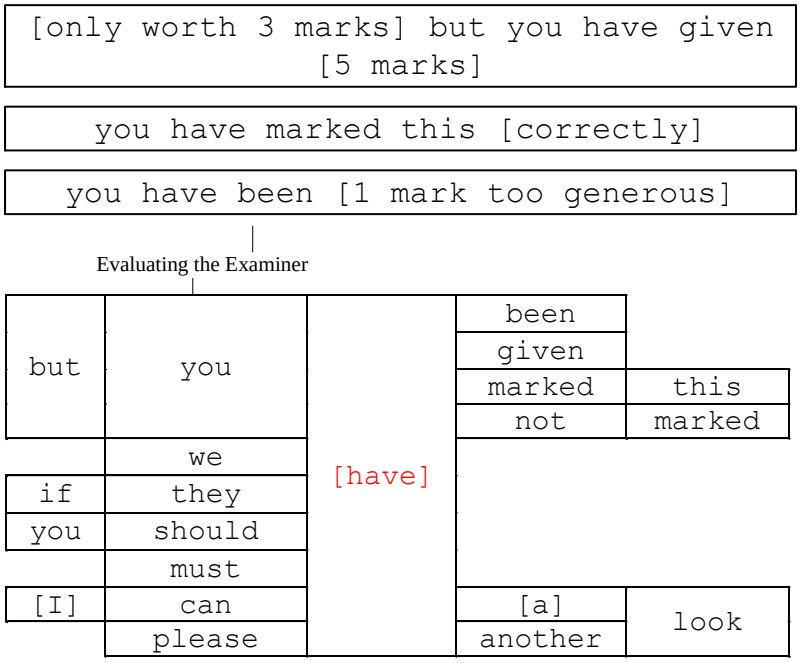

Figure 5.28: Evaluating

Finally, my analyses (outlined earlier) also showed that giving *Technical Information* was more common in Economics messages compared with the other subject areas (Table 5.9). Again, my theoretical framework does not adequately help to explain this finding so it requires additional consideration.

Having outlined my first SCDA theme, *Content*, I now move onto my second theme, *Time*.

5.2.2 Time

My analyses showed that some keywords and targeted search words were used significantly more in some quartiles rather than others. The table below (Table 5.14) shows the distribution of words pertaining to each type over the four discourse quartiles (for the sake of transparency, and owing to a lack of space in the chapter, I have presented the list of keywords for each quartile and their significance statistics in Appendix M).

In the earliest quartiles, there is a suggestion that there is more directive activity than in the latter quartiles (`gave, awarded, correct, mark, here`). In the final half of the discourse, there is a greater presence of terms that are potentially associated with exploratory

discourse (because, thought, if, think) and personalised connections with actions (you, your, me, my). This observation coheres with my theoretical framework and suggests that the examiners are more likely to indicate the location of marks in the earliest stages of the discourse, and then to engage with the reasoning for differences of perspective later in the marking discourse as trouble sources around specific marking issues appear. According to my theoretical framework this may represent the initial stages of the examiners exploring the extent of their shared common ground prior to carrying out extra remediation work once they realise that there are differences in perspectives to resolve through additional feedback.

Table 5.14: Keyword Use and Quartiles

Word type			Quartile			
TL	Exploratory	Reference	Q1	Q2	Q3	Q4
mark						
here		here				
correct						
only						
answer						
some						
gave						
awarded						
need						
please						
first						
was						
also						
rest						
tolerance						
etc						
seed	because	again				
	thought					
bit	if	they				
many	think	you				
		your				
		me				
		my				

5.2.3 Joint Intellectual Action

Having outlined the second of my SCDA themes, *Time*, I now move onto the third theme, *Joint Intellectual Action.* In this section I deal with the *Bridging* theme, where three codes

(being applied 993 times in total) accounted for 77% of my coding in this area. These codes were *External Reference*, *Historic Reference*, and *Offer Support*.

In my theoretical framework the concept of *Joint Intellectual Action* (JIA) describes indications in the discourse where participants display awareness of the actions of another participant. An implicit element of this concept is that one participant is thinking about how the other participant is thinking. This concept can sometimes be evidenced in evaluations where a TL comments on the examiner's performance to consider whether the thinking underlying that performance is legitimate. My analysis of the feedback suggested that the examiners accomplished these acts in a variety of ways, which I was able to cluster into four Bridging Themes.

5.2.3.1 The TL makes Links between Information Sources that they perceive to be Relevant to the Examiner (Bridging)

Bridging is an indication of where a participant brings together information in a feedback message. I take this to indicate that the participants have made a decision to include content that they consider to be appropriate given their assessment of the other's level of understanding. The concept also describes the way that language was used to directly bring together concepts or objects in the discourse. This included the use of communication to indicate consequential actions (e.g. linking a mark award to a preceding rationale [*Historic Bridge*]) or to orientate participants (e.g. to common areas in shared examination papers [*External Bridge*]). The example below (Figure 5.29) shows how the keyword [so] is used by TLs to link concepts found in the performance scripts with the consequences (in terms of marking outcomes).

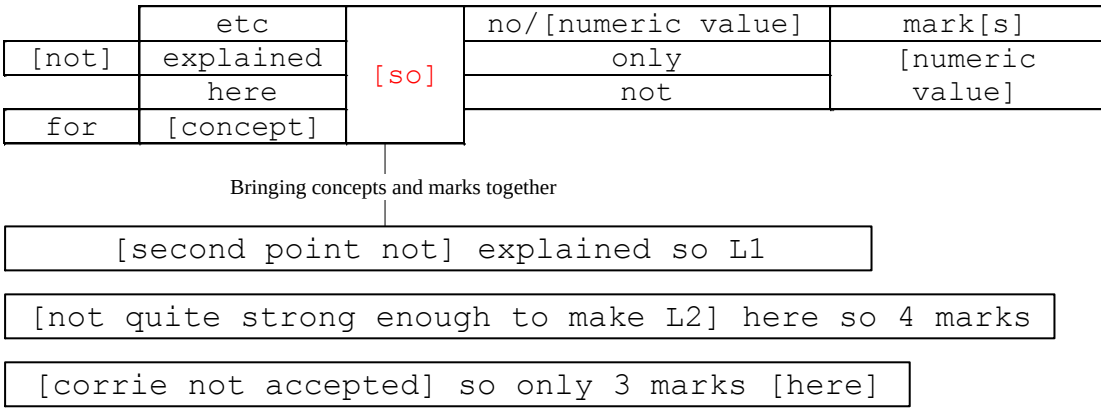

Figure 5.29: Bridging

The example below (Figure 5.30) shows how [put], a word significantly overused in the 1st/2nd Quartiles, is used by participants to support the development of shared focus. In the example below the TL uses the target word to draw an examiner's attention to the location of annotations which become the location for shared discourse.

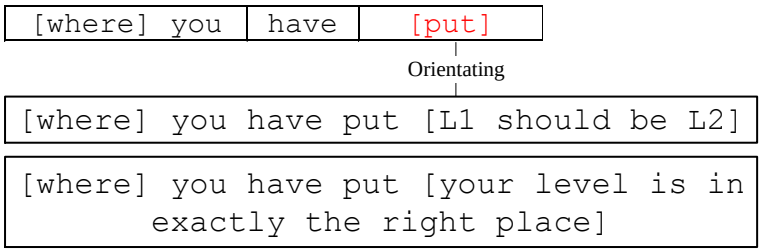

Figure 5.30: Shared Focus

As I stated earlier, an implicit element of JIA is that one participant is thinking about how the other participant is thinking. There is evidence in the data that TLs and examiners use questions in the process of eliciting and elaborating thinking. Whilst the role of questioning in institutional talk (e.g. classroom talk) is well documented (e.g. see Ingram & Elliott, 2016; Zemel & Koschmann, 2011), this is not the case in studies of examiner discourse.

5.2.3.2 Using Questions to Support the Clarification of Communication

The next Bridging Theme centred on how the participants used questions to clarify communication. Extract E8_30.5 18:06/30.5 1949 (Figure 5.31) is an episode of examiner question and a TL response. Here it appears that questions are used to focus the interaction. These allow the participants to stake out their position with regard to acceptable interpretation of the mark scheme. In line 005 [Thanks] is used in an anticipatory fashion to prompt a response, in keeping with turn taking norms in conversation (Fraser & Nolen, 2009). The examiner uses the question to elicit information from the TL, which acts to focus the direction of the discourse. The TL response is concise (line 006), affirming the examiner's explanation.

```
Ex: 30.5 18:06
001   I know it states in the mark scheme in brackets that it needs
002   specific reference to the diagram - just to check does this
003   therefore mean just stating price and qty increases is not
004   enough without linking to diagram?
005   Thanks

TL: 30.5 1949
006   Yes - they must link to the diagram.

[E8_30.5 18:06/30.5 1949]
```

Figure 5.31: Question Use

Questions are also used to invite participants to state their position of thinking, or to reflect on this position. It is possible that questions focus discourse in a less face-threatening way than more directive discourse because they invite the participant to reflect on the communication content. In extract E15 22.6 T 51-84 (Figure 5.32) the TL wants to remediate the examiner's marking so that they raise the score (lines 001-002) [you've only given it 5]. Through the discourse the TL changes their evaluation (line 032) [no, I think that one's about right really]. This shift of position appears to relate to the examiner having been able to state their own perspective of quality, which is aligned with the concepts of [depth] and [detail] (lines 018/025). The examiner uses questions to focus the discussion, with questions being perhaps less confrontational than statements.

```
001   TL    Now the essay there you've given 5, I think ::: 5, you've
002         only given it 5. Have a look at it ::::
003   EX    Right :: ok?
004   TL    Now ok :::
005   EX    Sorry, it keeps flicking all over the place on questions 5
006         and 6. You lose the front and then you, right
007   TL    Now the beginning bit is all rubbish, is all description
008   EX    Yeah
009   TL    So that's all ::: I would just stick "Seen" by something
010         like that, on the side.
011   EX    Yeah
012   TL    And on the ::: that's not credit worthy, that's not credit
013         worthy :: short term impacts included power cuts. Why?
014         There's no explanation.
015   EX    Yes, that's why I put the [arrow there, yeah]
016   TL                              [the] main impacts were economic
017         ::: household property, but there's no ins-::
018   EX    There's no depth to it is [there?]
019   TL                              [there's nothing] there
020   EX    Yeah
021   TL    And a dash, that's all description,
022   EX    Yeah
023   TL    A number of social impacts, such as death, loss of
024         property, no link to [water]
025   EX                           [There's no] detail again though is
026         there TL name?
027   TL    No, no. This is a good bit here, positive impacts for
028         example high water levels have meant that fishing, oh no,
029         that's rubbish because I don't know whether that's
030         right to be honest::: alternative :: right I think that
031         one's ::: more or less right. Perhaps 5 is a bit ::
032         uneconomic :: no, I think that  one's about right really
033   EX    Ok?

[E15 22.6 T 51-84]
```

Figure 5.32: Question Use and Face Management

In extract E11_529.1225/529.1736/529.1800_Q2 (Figure 5.33), which is a record of three interlinked written messages, the TL question (lines 006-007) prompts a direct reply in which the examiner acknowledges that they do not know the answer. This leads to them posing their own question to the TL (lines 014- 015) [I am assuming it is more standardisation scripts?], prompting a clarifying response from the TL. This discourse structure mirrors the initiation-response-feedback patterns observed in much classroom discourse (Mehan, 1979; Sinclair & Coulthard, 1975) and mirrors the function of

some of the questioning strategies observed in classroom discourse. Wegerif (2008) highlights how this form of questioning can be predicated on the dialectical notion of communication that underpins sociocultural learning theories that are inspired by Vygotsky (1986). Where questions are used to support a dialectic function, they are used to control and direct the learner towards an expert teacher's intended learning goal. This is demonstrated in the example below, where the primary focus of the TL question is to control the flow of information and to reinforce discipline on the examiner. This is done through asking the examiner to self-reflect, as part of the process of working out what the TL already knows.

TL: 529.1225
001 Hi *Examiner name*
002 Can you tell me why I asked you to wait and do these scripts
003 one at a time please? I am not being funny here I just want
004 you to understand what I am trying to do here so that you
005 understand and cooperate fully.
006 Also, as you did not wait, can you see what you are going to
007 have to do now?
008 *TL name*

Ex: 529.1736
009 Hi *TL name*,
010 I have not marked any more of these scripts today, other than
011 the two questions which I couldn't see the bottom of last night
012 which I thought you told me to go ahead and do in your last
013 email to me […] I am afraid that I don't know what you mean
014 about what I am going to have to do now, although I am assuming
015 it is more standardisation scripts? […]

TL: 529.1800
016 Hi *Examiner name*
017 Yes, if you had done one at a time it would have, hopefully,
018 have given you more information for the next one and so on.
019 That was why I wanted you to do that now and for the practice
020 scripts also. Well, what next? I do not want to give you the
021 second bunch of standardisation scripts. This was brought
022 in for the first time last year and it is arguable as to
023 whether that was very productive. However, I do think that I
024 will have to let you mark another batch of five of the live
025 marking, NOT YET!

[E11_529.1225/529.1736/529.1800_Q2]

Figure 5.33: Question Use and Clarification

It is perhaps not surprising that TLs employed question use as a strategy in their feedback, as questions have a recognised potential to encourage learning interaction. The ethnomethodological perspective which informs my theoretical framework suggests that the power of questions resides in their ability to elicit a response; i.e. they implicate a respondent into returning a response. This implication enables TL to use the feedback stage of the initiation-response-feedback structure to engage in scaffolding strategies because the question can furnish the TL with information about the examiner's knowledge state, that they can then use to plan the next communicative intervention. When seen from this perspective, the TL's feedback response is in effect a form of 'repair act' (Schegloff et al., 1977), as the TL seeks to impose cognitive order (Schegloff, 1992) in a particular discourse episode.

There are two potential mechanisms through which TL questions might encourage examiner learning. The first of these is the suggestion that questions can be effective at prompting higher order thinking and self-reflection in the recipient. Building on the ideas of Bakhtin (1981) and Vološinov (1973), questions can encourage the examiner to engage with their own thought position in relation to the dominant ideas in an area of thought, which are in this case represented by the TL's perspective (i.e. the 'authoritative discourses of a cultural canon', Maybin, 2001, p. 65). Cognitive development is therefore a product of the way that the question asks the examiner to engage with the dominant ideas and to explore their own thinking in relation to these. The second mechanism through which questions might encourage learning relates to what a TL does with the examiner response. Here, the question is a resource that encourages a learning interaction between the TL and the examiner, with the shared words (and their embedded ideas and concepts) being a resource for the participants to use as they strive to maintain the common ground with each other that can promote productive development (Edwards & Mercer, 1987; Horton & Keysar, 1996; Littleton & Mercer, 2013).

Analyses showed that TLs used feedback to invite examiners to either elaborate their own thinking or to expose their need for further information from the TL. An invitation for clarification represents a point where the participants expose the limits of their own knowledge of the thought position of the other on a social level. The example below (Figure 5.34) shows how the keyword [please] is used by TLs to invite the examiner to keep the discourse flowing if they lack sufficient information from the TL. This recognises that an outcome of such discourse flow is the potential for the fine tuning of communication to lead to the reduction of misunderstandings.

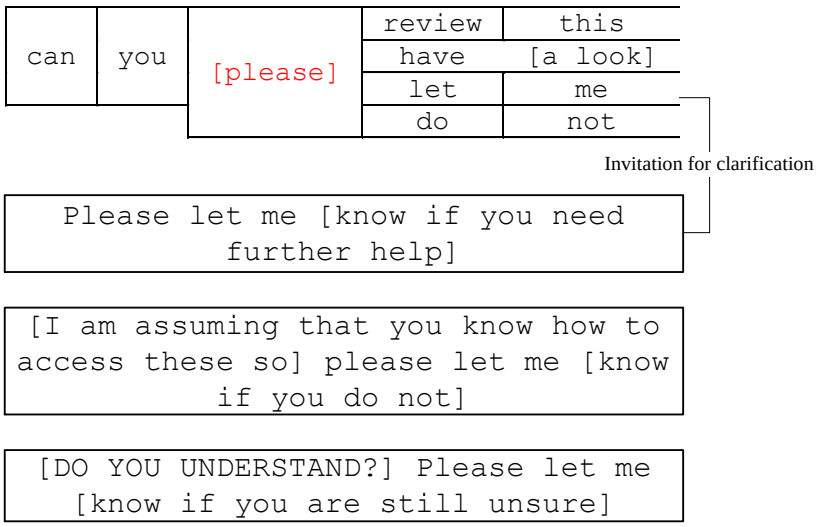

Figure 5.34: Politeness and Invitation

A feature of JIA is that participants work together through discourse to achieve a goal. Some extracts demonstrate how, despite the hierarchic nature of the discourse in general, the examiners were able to take an active role in the communication. Extract E2 a 617.0001 (Figure 5.35) shows how an examiner steers the focus of the communication, moving from a closure opportunity (line 021) [Right, ok] to go on to inform the TL about their practice. This is followed up with a direct question to ask whether the TL also does this (line 024), which leads to a confirmation by the TL. The examiner also indicates that they anticipate taking the lead in interaction, moving the focus of the interaction from a reflection on past practice to potential future interaction (line 027) [I'll be in touch]. This follows an indication by the TL that they are closing down the interaction (026) [but nice to talk to you].

```
010   TL    Oh good, fine, so you're up and running
011   EX    Sort of I wouldn't say [running]
012   TL                           [no] I know it's quite a marathon
013         I'm afraid this one this year
014   EX    But erm not much that seems very awkward:: not much
015   TL    No I think 7 can be a little interesting [but]
016   EX                                             [yes]
017   TL    But most candidates by the time they get there are quite
018         tired and so their answers aren't brilliant. There's a few
019         little bits and pieces there to watch out for. I'm sure
020         you'll cope
021   EX    Right ok thanks I'll erm just sort of cutting it up into
022         pieces to [to the wall]
023   TL              [laugh]
024   EX    Will you do that?
025   TL    I use a piece of board and pin it to boards so I can then
026         move it around a bit but yes (laugh) nice to talk to you
027   EX    Alright thanks I'll be in touch
028   TL    Thanks bye

[E2 a 617.0001]
```

Figure 5.35: Active Examiner Participation

In extract E9 25.5 T 169-177 (Figure 5.36) the examiner offers a personal rationale and reflection that supports the development of common knowledge that is used for further discourse. In the initial opening, the discourse shifts from a quantitative focus on misaligned marking to an explanation of the qualitative dimensions of the disagreement. The TL imposes authority through anticipating what the examiner is thinking (lines 001-002) [you've probably not given it L3 because], and gives a diagnosis based on examiner affirmation (line 007) [yeah]. The minimal gaps in discourse suggest that the achievement of agreement is the preference in this episode. As the communication develops, the sections build on each other and confirm that the previous meaning is received. This development centres on a common reference [Bands] (lines 001, 004, 009, 020). The TL dominates the discourse in terms of words expressed, using these words to sanction and refute the examiner at times, but the examiner also has a role in offering information to open up a potential for building common knowledge. This is demonstrated where the examiner adds information to the interaction (line 019) [i.e. the production and the matrix analyses] which gives the TL extra ground on which to refute the examiner's point.

001 **TL** Looking at the mark scheme at Level 3 Band 1, you've
002 probably not given it L3 because you've not seen that
003 first bullet point. That first bullet point isn't
004 required; it's just an example of what Band 1 could be.
005 So what they've not done, is they've not said *the cost of*
006 *production rises it shifts the supply curve* [but]
007 **EX** [yeah]
008 **TL** If you've got supply curve shift and price rise and
009 quantity fall, that's L3 Band 1, but this answer is
010 definitely not going to get above 10 because it doesn't go
011 on to do anything else, it doesn't relate back to market
012 failure. And even if it were to say, by the way, from
013 there, *and as quantity has fallen from market failure*
014 *therefore it's externalities are reduced* that wouldn't be
015 enough, it would have to be *so production has fallen which*
016 *solves overproduction, has reduced welfare loss*, it would
017 have to be something a bit more tangible
018 **EX** So basically Band 3 is the direct link into the question
019 i.e. the production and the matrix analyses
020 **TL** Yeah – that's the Band 2, and then until it gets into Band
021 2 it can't get any L4 marks
022 **EX** I sort of got that, you can work that out half way through
023 the essay because if they haven't done it half way then
024 they're unlikely to get it
025 **TL** Very unlikely, so then you end up just flicking through
025 and you see all this other stuff that's mentioned and it's
026 not really going to get any credit
027 **EX** And especially when they go off about all of these other
028 alternatives and then don't really do anything else with
029 it
030 **TL** Absolutely. So that's that script

[E9 25.5 T 169-177]

Figure 5.36: Preference for Agreement

In extract E3 CW b 620.0900 (Figure 5.37) the examiner steers the focus to try to better understand the TL's perspective. Eventually this examiner action leads to clarification from the TL. There are a number of pauses that act as dispreference markers in response to the TL's news that there are scripts to re-mark (lines 012/017/019). Persistent examiner questioning (lines 017/024) [`can you tell me`] which fail to exact a suitably clear response leads to a direct question (line 029) [`is it am I giving them too much`]. These questions shift the TL from deflecting the conversation from dealing with the issues immediately by giving a vague response (line 013) [`there's one I think 9 or 10 differences and one was about 7`], (line 018) [`in your workspace`], (line 020) [`sent you feedback`], and downgrades (line 014) [`nothing serious to worry about`], (lines 027-030) [`little bits`] [`just little bits`]. Eventually the TL acknowledges the source of the complexity, which is that the mark scheme has opaque elements that only make sense when looked at in relation to a particular script.

```
006   EX    Not been up too long
007   TL    [laugh] sorry>
008   EX    it's alright it's ok
009   TL    the morning after [laugh] erm I've sent a couple of
010         scripts back which I would just like you to re-amend and
011         send in and then I'll approve your marking>
012   EX    ::a couple of them to have a look at?
013   TL    Yeah there's one I think 9 or 10 differences and one was
014         about 7 and erm it's::: nothing serious to worry about
015         just have a look at them resubmit them please and I'll
016         approve you
017   EX    ::can you tell me which two they are?
018   TL    They're back in your workspace as we speak
019   EX    :::right erm
020   TL    I've sent you feedback as well, so I've sent you a couple
021         of messages. One with the feedback on those two and so you
022         can do those and another with feedback on the remaining
023         eight which were fine
024   EX    :::erm can you tell me which question it was so I've got
025         an idea and I can be thinking about it was it a mechanism
026         one o::r
027   TL    no it was a:ll through the paper there were little bits
028         and pieces where we differed
029   EX    is it am I giving them too much or not enough
030   TL    no no I think it was just little bits. The trouble is
029         there are some bits and pieces that are not actually on
030         the mark scheme specifically, there's sort of an implicit
031         deduction from what's on the mark scheme from what they
032         write you've got to make your own deductions whether
033         they've got the mark or not. So have a look at it first,
034         yeah have a look at this script have a look at my comments
035         and see where where the differences are

[E3 CW b 620.0900]
```

Figure 5.37: Examiner sets the Agenda

Having looked at question use as an indicator of JIA, the next Bridging Theme looks at how TLs referenced examiner thinking during their evaluative feedback communication.

5.2.3.3 Evaluating Examiner Performance

This theme describes the way that language was used by TLs to convey their evaluation of the examiner's marking performance. This language use is often collocated with the use of reference target words to outline the TL's understanding of the examiner's work. In this way, the TL is assuming a thought position that they believe is held by the examiner. The example below (Figure 5.38) shows how the keyword [which], that was overused in the 3rd/4th

feedback quartile, collocates with the reference keyword [you] and is used by TLs to convey their evaluation of the examiner's performance.

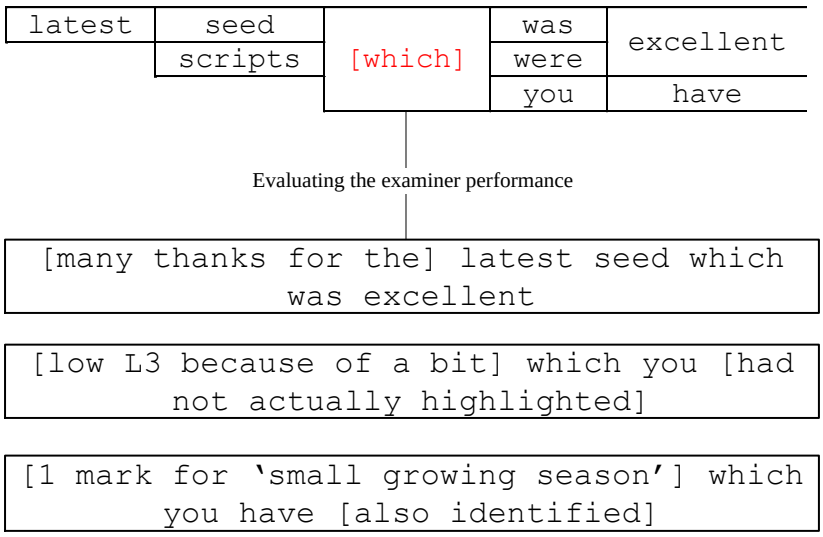

Figure 5.38: Evaluating Examiner Performance

The final Bridging Theme looks at how TLs referenced examiner thinking through indications of contrast.

5.2.3.4 The TL Contrasts their own Practice with that of the Examiner

This theme describes the way that language was used by TLs to contrast their own marking perspective with that of an examiner. This function tended to involve the use of the exploratory target word [but] to contrast the examiner's actions with the TL's perspective. This juxtaposition is based on the TL recognising that there is disagreement between their position and that of the examiner. It also allows the examiner to consider the basis for their disagreement with the TL (Figure 5.39).

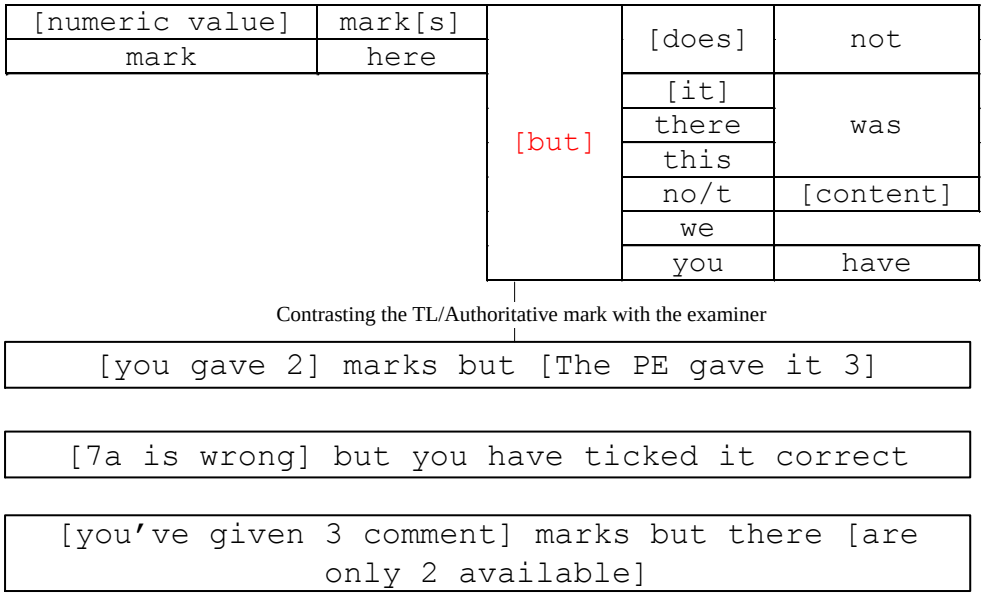

Figure 5.39: Contrasting with the Examiner

Having outlined the third of my SCDA themes, *JIA*, I now move onto the final theme, *Impact.*

5.2.4 Impact

In this section I deal with the *Impact* theme, and this included three codes (being applied 691 times in total) which accounted for 100% of my coding in this area. These codes were *Review, Resolution,* and *Reification*.

An indication that agreement is reached through discourse is evidence that resolution has occurred. This also indicates that the thinking of the participants has reached a degree of working alignment so that they can continue with their work. The fine detailed analyses that I carried out using ASCDA methodology enabled me to gather evidence of how examiners attained resolution and alignment. In extract E15 22.6 T 199-217 (Figure 5.40) the TL seeks active agreement from an examiner following them receiving a limited response to a request for information. In line 001 the TL invites the examiner to explain but then does not engage with the reasoning given. From this point, the examiner accepts TL decisions without offering information in return (lines 13, 15, 20, 27). It appears that the TL senses that the examiner is accepting judgements uncritically and uses another question to interrogate further (line 008) [Do you agree with me with that? No?]. This then leads to a TL elaboration move to ensure that an important conceptual point is reiterated (lines 11-26). Although the TL dominates the interaction, there is an apparent concern that the examiner is

not engaging in active agreement, and that this engagement is important for fine tuning the way that the feedback discourse develops.

```
001   TL    And 5, do you think that's too generous?
002   EX    Well. I was trying to be generous because they've actually
003         attempted the conclusion [??
004   TL                            [Yeah, yeah, yeah, yeah]
005   EX                             ??] which was absolutely [atrocious
006   TL                                                      [4] give it 4.
007   EX    Right, ok boss. Done.
008   TL    Do you agree with me with that? [No?]
009   EX                                    [Totally] totally and
010               utterly
011   TL    Ok. So what you're looking for then, are needs, there's
012         got to be needs,
013   EX    Yeah
014   TL    And the bigger the variety the better
015   EX    Yeah
016   TL    And :: any, erm :: natural process at all :: erm :: now
017         what's hovering at the b- :: yeah, and it can be
018         weathering as well, it doesn't have to be erosion, it can
019         be weathering or erosion, mass wasting, anything like that
020   EX    Yeah
021   TL    It can be long shore drift : [because
022   EX                                 [right]
023   TL                                          that] takes the
024         beach away so that's a natural process and in discussion
025         if they say, no the beach isn't there to protect the cliff
026         then that'll be the icing on the cake.
027   EX    Yeah
028   TL    Ok. Let's have a look at 7 and 8.

[E15 22.6 T 199-217]
```

Figure 5.40: Seeking Active Agreement

In extract E1 618.0001 (Figure 5.41) information that is shared in the early stage of a new examiner/TL interaction influences the approach taken by the TL in the later stages. The concerns of the examiner over the time needed for their work is cued in line 020 [`just getting the first five done`], this leads to the TL accentuating that the examiner needs to take time (lines 022/024) [`time`] [`l:ong`] [`marathon`]. Later in the conversation, the reason for the time concern is evident; enquiry about their familiarity with the marking technology leads the examiner to divulge that they are marking another examination unit

(lines 027). The TL appears concerned about this, indicated by the dispreference gap on line 028 [r::ight excellent].

```
016   EX    It'll come to me first d'you think
017   TL    It should come today through email as well as being on the
018         RM site, as soon as I hear I'll drop you an email and let
019         you, know, is that ok?
020   EX    It's just that getting the first five done on time,
021         y'know?
022   TL    Oh I promise you, you need to take your time with these
023         because it's um I think the mark scheme is about 29 or 30
024         pages l:ong, it's a bit of a marathon
025   EX    Ok then, thanks for that?
026   TL    Now have you marked on scoris before?
027   EX    Yeah I do F324
028   TL    Oh r::ight [excellent]
029   EX                 [So] this year I wanted to do y'know a
030         different one, a little bit of something else
031   TL    Excellent indeed. So you're familiar with classic scoris
032   EX    I'm familiar with how to go about all that [yes]
033   TL                                              [Excellent, good]
034   EX    I just want to get the F322 out of the way
035   TL    Yes, indeed
036   EX    Yes before I start the F324 one

[E1 618.0001]
```

Figure 5.41: Adaptation of Discourse to Reach Resolution

Another indication of resolution is where the participants actively mirror each other's language use. This phenomenon may represent the development of a common language, with participants coming to use a more refined vocabulary with shared understandings around the use of specific terms.

In extract E11_602.1708/602.1849_Q3 (Figure 5.42) the examiner replicates the language of the TL (lines 001/002) [pleased] [keep this consistency] and (lines 005/006) [pleased] [keep it up]. This repetition may have a number of functions. According to Bowe & Martin (2007), repetition can reassure the initiator that the recipient is attending to their words, or it can suggest that the recipient is understanding and/or agreeing with the initiator. Repetition may also imply the development of a shared language, with recycled words being a resource for common ground building (Schwarz, Kolikant, & Mishenkina, 2012). The existence of repetition has been observed by Bowe & Martin (2007) in other

professional learning contexts, and coheres with observations based on accommodation theory (Giles, Coupland, & Coupland, 1991).

```
TL: 602.1708
001   […] You can be very pleased with this beginning, the trick is
002   now to keep this consistency!
003   TL name

Ex: 602.1849
004   Hi TL name,
005   Thank you, I am very pleased (and relieved)- as you say, I now
006   need to keep it up! […]

[E11_602.1708/602.1849_Q3]
```

Figure 5.42: Linguistic Repetition

In extract E8_529.1947/529.2050/529.2100_Q3 (Figure 5.43) the TL prompts the examiner to explain their thinking (line 009) [`why did you keep this at 10?`]. This leads to the examiner introducing the concept of the convincing argument (line 015). The TL then fine tunes the feedback message by repeating the reference to [`convincing`] whilst shifting the argument to the importance of the presence of certain [`trigger phrases`] (line 020) in the performance.

```
TL: 529.1947
001   Hi Examiner name,
002   This is quite a bit out and both this and the last script have
003   been awarded the wrong level in the essay. Can you please
004   review?
005   1 - maize production or leisure activities = application of
006   unlimited wants.
007   7 - they have explained rival and said it is non-rival, so they
008   can't get that explanation mark.
009   8 - why did you keep this at 10?
010   Thanks,
011   TL name

Ex: 529.2050
012   Hi TL name
013   For Q8 I felt that they explained subsidies, but didn't really
014   make it clear how market failure was corrected by them. I
015   wasn't convinced, but obviously should have been.
016   Examiner name

TL: 529.2100
017   Hi Examiner name,
018   As soon as they say it solves the underconsumption problem that
019   is enough - not convincing perhaps, but gets lucky with putting
020   in the trigger phrase!
021   TL name

[E8_529.1947/529.2050/529.2100_Q3]
```

Figure 5.43: Explanation and Fine Tuning

5.3 Summary of My Findings for Research Question 1

In this chapter I have used the four areas of interest that arise from my theoretical framework to consider the question: 'What are the characteristics of examiner feedback?' To do this I looked at *Content* (at levels of interaction and transaction), *Time, Joint Intellectual Action*, and *Impact*.

My theoretical framework anticipates that feedback would act as a repair mechanism and centre on a trouble source (which is the existence of a weak mutual common ground). This framework also suggests that common ground establishment, which includes work related to the construction of an Intermental Development Zone (IDZ), is a component of learning discourse, and that this work would be most pronounced at the earliest stages of feedback interaction.

My analyses offer some evidence to support the idea that the more subjective, levels-based marked items are a trouble source, with more feedback being given for those examination papers that include the most subjective items. This trouble source contributes to feedback content being heavily infused with disagreement. The prevalence of feedback being used to locate where marks should be given (without explaining why they should be given) in the least subjective examination paper also suggested that item type was an influence on feedback complexity.

My analyses also offer some support for the notion that feedback was functioning to align perspectives and diminish common ground weaknesses with most being communicated to new and unfamiliar examiners, and most being found at the initial stages of communication, where learning needs were the greatest and shared common ground the weakest.

As well as the consideration of how feedback messages change over the feedback period (i.e. across message development), my analyses also offer an insight into decisions about where information is placed in individual messages (i.e. 'within message' chronology). An interesting observation from my analysis was that the construction of message opening and closing structures were purpose driven and aimed at creating an effect. Message openings were the location of interpersonal distancing strategies that included indications of politeness. Positive interpersonal relationships help to reduce the social or subjective distance between interaction participants (Cramton, 2001; Wilson et al., 2008), and it is possible that the use of politeness at the start and end of feedback messages performed a relationship management function. This was especially the case for the most subjective items (Geography and Economics), for new and unfamiliar examiners, and particularly given the prominence of negative information content in the feedback messages as a whole.

My theoretical framework anticipates this issue, drawing on the work of Goffman (1967), who argues that the presence of negative information in social interaction embroils issues of face management, and that politeness devices and 'redressive facework' can be used to reduce friction (e.g. Culpeper & Haugh, 2014; Lakoff, 1979). I argue that the prevalence of negative discourse has implications for professional face management, and that the examiners would be expected to engage in mitigation strategies (i.e. articulation work) to maintain positive professional relationships.

According to my theoretical framework this articulation work helps to shift the discourse from disputational to cumulative discourse. This cumulative discourse is important as it

enables the examiners to continue to engage with differences in perspectives and to develop their common ground through bridging perspectives (e.g. a shared focus on resources).

My qualitative analyses showed that articulation work involved a number of strategies that appeared to increase the symmetry of status between the examiners as the TL actively sought to engage the examiner in resolution. These strategies included shifting the focus for disagreement, underplaying differences, using negative politeness, and structuring information around the openings of messages to create a specific effect.

Finally, there were two findings that I could not explain through my theoretical framework. I was not able to anticipate the use of *Directive* and *Technical Information* in Economics feedback in particular, and so this may need additional refinement of my theoretical framework.

6. Results: What are the Characteristics of Effective Examiner Feedback?

In this chapter I explore the features of feedback communication that may be considered to be 'effective'. According to my theoretical framework, effectiveness is signified by the alignment of TLs' and examiners' perspectives, and this is evidenced by the attainment of resolution in discourse.

My approach to the analysis of effectiveness in this chapter has two parts, one that is theoretically driven and another that is problem-oriented and empirically driven. By adopting these two approaches, I seek to dissect the feedback data (so as to get an in depth view of particular communication features) whilst also constructing a broad overview of the data (through integrating my observations with theory to create a coherent and perhaps more generalizable picture of the communication). In this way I also seek to overcome the criticism that research can lack engagement with 'the generalizations that have built up over the years in any one discipline of the behavioural sciences', whilst also failing to qualitatively explore really important issues (Lawrence, 1992, pp. 140-142).

For the first part of the analysis I relate the features of feedback (that I identified in Chapter 5) to research literature that I have organised into a 10-factor Integrated Analytical Framework (IAF) (Johnson, 2016c). I have outlined the IAF in detail in Chapter 2, but I will also give a brief overview here to frame the following discussion. The IAF emerged from my analysis of feedback research literature, and it integrates factors from a broad array of research studies. The IAF comprises 10 factors with attached dimensions (Table 6.1).

Table 6.1: The Integrated Analytical Framework (IAF)

	Factor	Sub-component	Qualities		
Core	1 Language Use	a) Clarity of communication b) Discourse characteristics	Clarity Cohesion	←→ ←→	Vagueness Distancing
	2 Content	a) Detail b) Quantity	General Restricted	←→ ←→	Specific Elaborated
	3 Timing	a) Temporality b) Frequency	Immediate Often	←→ ←→	Delayed Infrequent
	4 Form	a) Mode b) Interaction condition	Oral Collocated	←→ ←→	Written Remote
Socio-cultural	5 Source	a) Number b) Trust	Singular Respect	←→ ←→	Multiple Distrust
	6 Emotion	a) Evaluation	Positive	←→	Negative
	7 Recipient	a) Feedback seeking b) Confidence c) Locus of control	Avoidance Secure Self	←→ ←→ ←→	Seeking Insecure Other
	8 Knowledge	a) Codification	Codified	←→	Tacit
	9 Feedback Giving	a) Explicitness	Explicit	←→	Implicit
	10 Pedagogy	a) Model of learning	Transmission	←→	Construction

The framework is an organising device that can be used to consider the individual elements that influence feedback impact. The dimensions are either variables that can be manipulated by the feedback giver (e.g. language use) or features that can be taken into account when considering the impact of feedback (e.g. recipient characteristics). I argue in Chapter 3 that dimensions of the IAF can support the development of an Intermental Development Zone (IDZ) (Mercer, 2002, 2008b; Mercer & Littleton, 2007). The development of the IDZ is central to feedback being effective because communication that opens up this zone enables participants to establish common ground with each other and this is the where alternative perspectives can be brought to resolution. The IAF and the list of features of effective feedback that it proposes allow me to illustrate these features 'in action', but it is actually the demonstration of whether alternative perspectives can be brought to resolution through the use of these features that is the key indicator of effectiveness.

The first section of this chapter (6.1) reports on the theoretically grounded phase of the analysis. This section uses theory to organise a description of the feedback features that either support or undermine convergence and the building and maintenance of common ground. To do this, I draw on the interview data from a sample of participants who self-reflect on their intentions behind, and reactions to, particular feedback messages.

The second section of this chapter (6.2) is oriented to the problem of how participants attain convergence through feedback. The section is organised around a series of cases that

exemplify how participants engage in extended episodes of feedback to reach convergence. In this way, the feedback centres on a shared problem (an examination question) and the cases illustrate how divergence is diminished through the participants' actions.

Finally, taken as a whole, the two sections of the chapter are used to inform the thesis outcomes and recommendations (Chapter 7) which generate lessons for Team Leader [TL] practice.

6.1 Effectiveness and the IAF

In the previous chapter I outlined how my initial analyses supported my theoretical framework which suggested that feedback centred on specific trouble sources that contributed to feedback content being heavily infused with disagreement. This analysis also showed how feedback was used to diminish common ground weaknesses (through the transaction of important information), whilst also involving the use of distancing (politeness) strategies. My theoretical framework suggests that this use of distancing is a form of articulation work that helps to shift the discourse from a disputational to a cumulative discourse and which increases the symmetry of status between the examiners.

In this section I explore the examiner SR interview data with a view to seeing whether the factors that comprise my IAF (which was based on extensive literature review) adequately explain examiner feedback effectiveness. As I presented in Chapter 2, the principle findings from the literature indicated that effective feedback involves:

- Discourse that seeks to support the construction of learning (*Pedagogy* factor), a participative dimension, and an active recipient role (*Recipient* factor: Feedback Seeking; Locus of Control). These findings recognise the important role of learner self-regulation in the learning process;
- Giving feedback immediate to task completion (*Timing* factor: Temporality);
- Implicit elements that centre on relationship management – which also implicate participant familiarity levels (*Feedback Giving* factor), alongside explicit elements that seek to codify important information (*Knowledge* factor). These dual purposes represent a key component of feedback complexity;
- Clear and specific language use (*Language Use* factor: Clarity of Communication; *Content* factor: Detail);
- Politeness and face management (*Language Use* factor: Distancing);
- Collocated interaction (*Mode* factor: Interaction Condition);

- The recipient having trust in the feedback giver (*Source* factor: Trust);
- Positive information feedback (*Emotion* factor).

My theoretical framework also suggested that hierarchic and remote communication arrangements presented a challenge to these elements of effectiveness.

In the next section I take each component of this framework and present a sample of the feedback and interview data (analysed using ASCDA methodology). The presented data seek to illustrate the key points of the analysis, systematically presenting how effective feedback can be realised in practice, but have also been chosen judiciously because of the constraints of the thesis word count. A more comprehensive presentation of data is included as Appendix N. It is important to point out that I am not claiming that the inclusion of these feedback features will necessarily increase the effectiveness, I am claiming that my analysis brings out the affordances and constraints that are evident with each factor.

Finally, I reflect on the interrelations and the contextual challenges suggested by my theoretical framework in the summary section at the end of the chapter.

6.1.1 Discourse That Seeks To Support the Construction of Learning

The analyses supported the idea that feedback communication is a site where differing personal perspectives can come together and be used as a resource for common ground building (Espasa & Meneses, 2010). The TL and examiner interview extracts below indicate how participant dialogue has an important role in the learning process. Dialogue allows participants to indicate areas of insecure understanding, to see a problem from another perspective (Edwards, 2012), and this opens up opportunities for fruitful interactions that allow the TL to reinforce those ideas that they consider to be most important (Whitelock et al., 2003).

> *[The examiner is] lovely, he asks a lot of questions and he's always messaging me.*
>
> *[Y2 TL2 [E24E25E26] SR interview]*
>
> *[Standardisation] is where you ask questions or if you see something that might get a mark or it gets a mark but you were unsure about it, you just sort of clarify and just there and then you can iron out the mark scheme.*

[Y1 E8 SR interview*]*

Analysis highlighted TLs' judicious use of information to encourage examiner development. This activity has parallels with scaffolding literature (e.g. Rojas-Drummond, Torreblanca, Pedraza, Vélez, & Guzmán, 2013), but also suggests some particularity. In the interview extract below, a TL explains the reasoning for a short feedback message:

> 5b(iii) Not propan-2-ene.
> [E23 M1 19.6.1045]

You can't give them too much, you would like them to go back into the script.

[Y2 TL1 [E19E21E23] SR Interview 128]

Here the TL alludes to the way that the provision of partial information encourages the examiner to actively engage in their learning construction. This links with the work of Hatzipanagos & Warburton (2009) who argue that feedback aids learning where it is part of a dialogue that supports shared conceptualisations of learning goals and allows learners to take more responsibility for learning. This suggests that, counter to the proposal that learner self-reflection should precede feedback (Harms & Roebuck, 2010), feedback is an important initial spur to examiner self-reflection through highlighting the existence of dissonance between the TL and the examiner. According to Recio Saucedo et al. (2013), such dissonance can instigate learner self-reflection.

Analyses also suggested that superficial learning occurred where feedback failed to encourage the active involvement of the examiner in sense making. Whereas elaborated feedback was considered by some examiners as a way of better understanding the philosophy of the Exam Board, so that they could better understand mark schemes, some examiners suggested that transmission-style feedback discouraged such activity. In their reflection on a piece of feedback, one examiner suggests that it is unimportant for them to understand the rationale behind the marking decision:

> 8a Must be a formula for water and the structure is interesting.
> [E5 M5 29.6.14]

8a is quite interesting because in the mark scheme it does say 'For E and F allow H2O/water'. And so I was giving a mark at that point if they put water. Now that is there in the mark scheme, however it would seem to be it must be a formula for water. So I start marking it to that point. Now in the previous year I probably would have gone back and queried that. But these days I don't bother because there is no point. If you are saying it needs to be the formula then I will just mark to that.

[Y1 E5 SR interview 158]

A consequence of this shift in engagement is that the examiner assumes a more technical-functional, rather than an active-reflective role. Building on Butler Shay (2004) this represents a weak sense of professionalism, as the value system that underpins the assessment process remains the property of the more powerful members of the marking hierarchy. This also has potential implications for the examiners' sense of professional identity and fulfilment. Interaction influences professional learning as it can encourage identity formation (Raz & Fadlon, 2006). According to Ren, Kraut, & Kiesler (2007) group identity grows through shared communication, with shared task work and social interaction increasing bonding. An aspect of this bonding also involves the sharing of personal information through self-disclosure in interaction.

Analyses suggested that TL feedback that elaborated the reasoning for decisions would help to articulate tacit and non-tacit elements of professional knowledge. In the telephone feedback extract below the TL articulates a key phrase [010 'so what'] that does not appear in the mark scheme, but is used to discriminate between lower level 'description' performances and higher level 'explanation' performances. This extract appears to be an instance of what Barton & Wolery (2007) call correcting and guiding knowledge in context. This feedback helps the recipient to learn through making sense of what Gasson (2005) terms 'the context-specific nature' of the knowledge, whilst also modelling ways of thinking and doing (e.g. Schön, 1983).

```
001   TL:   Yeah, you've been giving three instead of
002         one, and two instead of one, [and]
003   EX:                                [oh dear]>
004   TL:   six instead of five> but it's so nebulous I
005         I can't really, I think you've just got to
006         ::
007   EX:   I'm just being a pain in the butt aren't I,
008         be honest? [(laugh)]
009   TL:              [no] you've just got to keep
010         saying 'so what'
[E15 22.6 T]
```

According to Gasson (2005), codifying is a purpose of feedback. The articulation of key terms also helps to ensure that they are legitimised and not considered by learners to be suspicious (Stevenson, 2001). The importance of codifying knowledge through rationalisation is explained by an examiner when they reflect on the feedback extract below:

```
We didn't give the mark here because they didn't even
link heating to biofuels - so overall we assessed the
answer as not being relevant to the question set.
[E24 M3 2.6.1922]
```

Here he's given me the rationale and for me, as I say, it's that patience in explaining the rationale that's so important to me, because now I think, well yes I can apply this. I get it. I understand your philosophy now and I know why the [exam] Board wants this, so that's what I'm going to do. It's very difficult to line up with something that you don't understand, isn't it?

[Y2 E24 SR interview 166]

Literature suggests that personal attributes influence the reception of feedback (Ajjawi & Boud, 2017). Although I did not probe examiners' self-perceptions of feedback receptivity, my analyses suggest that some elements of the examiner context could induce anxiety prior to receiving feedback. My analyses suggested that features such as a lack of rich interaction at the initial marking stages can contribute to anxieties that can manifest in low levels of professional self-confidence. This sentiment is conveyed by an examiner in the interview excerpt below:

At the beginning [the TL] said 'Oh you're obviously very experienced', which kind of is quite a nice feeling for somebody to say, but I wondered in my mind 'Oh

does that mean that he's thinking I'm going to be fine', you know, kind of psychologically. Which was kind of a test at the beginning as an examiner because you're thinking 'Can I do it? Can I get through standardisation? Am I actually any good at marking?'

[Y2 E26 SR interview 111]

A TL also displayed the degree to which they were attuned to potential examiner anxiety when they discussed why they omitted communicating additional negative information to a specific examiner. The extract also alludes to how social isolation and communicating through media weak channels may exacerbate this effect:

```
The rest was fine.
[E19 M4 24.6]
```

There may have been one [question] in there which was a bit iffy, but I wouldn't in this instance bring it up, because it could be a BOD [benefit of the doubt]. And so therefore, it's only going to make this particular examiner more insecure than she is already. [The message] does have an effect, because you're sat in a room like this, on your own. It's very impersonal.

[Y2 TL1 [E19E21E23] SR Interview 740]

It was clear from the interview data that examiners and TLs considered that being a professional examiner meant that they possessed the locus of control for learning. This was often expressed in terms of the examiner being responsible for self-reflection as a result of the feedback that they had received. For Ajjawi, Schofield, McAleer, & Walker (2013) this is a critical feature that influences the effectiveness of feedback.

Analyses supported the idea that feedback communication that contained elaborated reasoning supported common ground building. This is demonstrated in the examiner reflection on the feedback extract below:

```
1a(iii) we gave a mark here.
[E5 M1 24.6.14]
```

So that is probably the least helpful kind of feedback to get back. What [the TL] should maybe do is to go back to the script and then look at it from there and say

'So why have you given him the mark there?' The temptation is to ignore that a bit really and think 'Oh well you know…' because again you are on the timeframe and you just think it's probably one of those odd little things. Actually kind of crack on with the next lot.

[Y1 E5 SR interview 57]

In their reflection, the examiner highlights the way that a lack of elaboration poses a specific risk to their alignment with the TL. The examiner suggests that a lack of detailed content outlining the TL's thinking, and therefore perceived as not articulating with the examiner's current knowledge state, left the examiner disengaged from any process of self-reflection, particularly in the time pressed marking context. It seems that a lack of elaboration limits the potential for the feedback to perform a feedforward function that affords future examiner self-regulation (Dannels & Martin, 2008; Furnborough & Truman, 2009; Prins et al., 2006).

Elaborated content tended to focus on important elements of misalignment that required remediation. For example, the feedback extract and the TL reflection below illustrate how this TL re-shaped the content of feedback in order to make it meaningful to the recipient:

```
7b No marks here - there is no mark for the IR
absorption alone here - it is in 6c where this occurs
and M can only score as an ecf when L is a ketodiol or
an ald-diol, as per Additional Guidance.
[E19 M2 19.6.2047]
```

I've basically reiterated what's in the additional guidance, but in a slightly different way.

[Y2 TL1 [E19E21E23] SR Interview 570]

In a subsequent interview the examiner who received this tailored messaging reflects on the positive effect that it has on their understanding:

[It's the rephrasing that helps, that sort of iteration?] Yes, yes, exactly, directly. Because it won't, it will not come again, I've understood it, let's do another way, another language, another way of saying the same thing.

[Y2 E19 SR interview 251]

This TL rephrasing activity in effect represented a re-coding of the message so that the examiner did not need to perform this challenging function (Knoch et al., 2007).

6.1.2 Giving Feedback Immediate to Task Completion

Analyses suggested that alignment was maintained through reassuring and frequent communication, which contributed to an on-going and cumulative interaction. In this way, the communication of frequent and immediate feedback supported common ground building. This coheres with previous analyses of feedback preferences in examiner contexts (Johnson, 2016a; Johnson & Black, 2012a), and in other professional contexts (e.g. Archer, 2010).

In addition, the data illustrates the challenge that communicating in a fast-moving marking environment places on TLs. Sometimes the ambition of providing adequate and timely support to examiners runs counter to the perceived demands of examiners. Here the concepts of time and content quality interact, with perceived over-communication interfering with task completion, in line with observations by Dennen, Aubteen Darabi, & Smith (2007). For example, the feedback extract and interview reflection below illustrates how time pressure steers the examiner towards a demand for restricted amount of feedback:

Examiner 29 to TL: For question 4 a(ii) if the candidate writes salt pan for a(i) and then explains how its shape has been formed using water accurately, is it possible for them to get full marks for a(ii)?

TL to Examiner 29: No because the landform is not found in this landscape. If we can we avoid double penalties but this time the PE decided not to accept these answers.
[E29 M5 25.5 15]

> *[I] don't need long blurb. If the answer can be yes or no then, you know … especially at this point.*
>
> *[Y2 E29 SR interview 53]*

The way that time pressure undermines the ambitions of feedback givers to offer elaborated, extensive feedback has also been noted elsewhere (Chetwynd & Dobbyn, 2011).

The apparent contradiction between the two previous points (i.e. that examiners may prefer both elaborated and restricted feedback detail) may be explained by the literature that suggests that elaboration and restriction relates to the recipient's learning state. For example,

elaborated feedback, which includes verification and guidance has been found to be effective for learning outcomes (Smits et al., 2008) and improving task comprehension (Murphy, 2010), but also risks 'over-communication' that can lead to recipients developing a negative self-concept of their abilities (Ackerman & Gross, 2010).

TLs and examiners also reflected on the cognitive demands that related to feedback immediacy. It was common for participants to suggest that delayed feedback was difficult to incorporate into their thinking during the learning phase. This reinforces points raised by Münzer & Holmer (2009) and Winter & McGhie-Richmond (2005).

An assumption in the common ground literature is that communication forms the basis of an on-going interaction that leads the participants to construct a base of common foundational knowledge. This perspective is conveyed by one examiner who stated:

> *It's kind of a cumulative effect … which is why it makes sense.*
>
> *[Y2 E29 SR interview 97]*

The assumption that examiners accrue experience in interpreting common terminology was challenged by the data. Examiners explained that the meanings of common terms could change from one marking session (year) to the next:

> *I think in previous, previous years when we've used [the annotation] EE it's been to show an extended judgement and I think this time, well this time they're using EE to clarify that that is an evaluation.*
>
> *[Y2 E26 SR interview 50]*

This means that the concepts that are the focus of examiner learning can be fluid and therefore require support at the start of each marking session. The shifting nature of language reflects the constructivist notion that communities generate and refine conceptual meaning through their language use. This builds on Wittgenstein's (2009) idea that meaning is 'use in a community' and is contingent on participants' mutual actions at a particular point of time.

6.1.3 Implicit Elements that Centre on Relationship Management

It is recognised that the need for feedback givers to make decisions when composing feedback is an element of feedback-giving complex activity (Ypsilandis, 2002). For the TLs

in this study, this complexity included making decisions about whether to give feedback at all. This sentiment is expressed by a TL:

> *We were generous on occasions because there are some [marks] that are a bit iffy, but then you go with the examiner. It depends on the individual. Again that is not a blanket statement because you have to know the people involved and their experience of marking.*
>
> *[Y1 TL1 [E2E5] SR interview 75]*

This decision-making links clearly to the ontological roots of the marking activity (Johnston, 2004), which affects the confidence with which a TL can definitely identify an examiner's judgement as being erroneous. This decision-making also implicates features of the TL's familiarity with and their knowledge of the experience level of the examiner.

Interaction affords an opportunity for participants to recognise and explore each other's perspective, and this is a prerequisite for common ground building. It is argued that effective feedback giving includes being able to look at feedback from the reader's point of view (Davis & Foster, 2002). This coheres with one of the three elements of communicative competence (Canale & Swain, 1980), which is an ability to perceive the needs of other participants. This element of mutual recognition was evident in episodes of feedback where the TL or the examiner overtly referenced each other's position, for example:

```
… so the answer was considered to lack appropriate
application. I can see the justification for awarding
the mark you have though and think it could be justified
in live marking.
[E11 29.5.1057]
```

> *[The TL] is basically saying it's an easy mistake to make and 'Yes, I can see why you might think…' I guess, [the TL] is reading my mind a bit there.*
>
> *[Y1 E11 SR interview 105]*

Analyses of the data imply that a TLs' self-understanding (e.g. recall of their own learning process) can be a model of how to best support examiners through feedback. This is reflected on by one TL during an interview:

[You said that you learnt about giving feedback from looking at it from the other side?] So, I felt like I didn't get given enough feedback, and so that left me feeling that by the end of the marking process I was not convinced that I understood the mark scheme fully.

[Y1 TL2 [E8E11E12] SR interview 132]

This reflection demonstrates how the TL uses self-reflection to overcome the concerns, expressed in the literature, that individual ego can be a poor basis on which to base a perception of others (Derks & Bakker, 2010), and that it can be a hindrance to bridging knowledge across a community (Boland & Tenkasi, 1995).

It is recognised that feedback-giving is a complex task that is under theorised (Archer, 2010; Evans & Butler, 1992). My analyses suggest that one element of this complexity relates to the nature of the decision-making that a feedback giver has to consider when constructing feedback. Besides considerations of core feedback features such as length, quantity and time (Ypsilandis, 2002), feedback givers also need to consider the situated needs of the recipient.

It is acknowledged that the attempt to satisfy dual functions in messages is a barrier to the effectiveness of feedback (Pryor & Woodward-Kron, 2014). Higgins et al. (2001) note that managing social relations and also delivering honest performance feedback represents such a dual function. There is evidence in my data of TLs being aware of covering a variety of purposes in their feedback communication, for example through the use of polite message openings:

```
Morning Examiner name
[E21 M1 23.6]
```

This is continuing a dialogue, so I've [just] spoken to [the examiner]. I think also it allows [the examiner] to feel involved.

[Y2 TL1 [E19E21E23] SR Interview 350]

In addition, analyses also show how feedback is used to give general marking principles, with these sometimes being interspersed among other, more specific pieces of performance feedback.

```
This script is almost as bad with 15 positive marks
over, at least they are both positive and we are not
looking at swinging from positive to negative here.

3 b  Very muddled answer, keeps referring to erosion,
not given any credit, 0 marks not 3
[E18 3.6.1055]
```

I said 'At least both of these scripts are positive, and we're not looking at swinging', so we have got some consistency. I am just trying to make it a little bit less [negative] but make the point that consistency is the key.

[Y1 TL3 [E16E18] SR Interview 168]

6.1.4 Clear and Specific Language Use

Analyses supported the idea that clear articulation and cogent language use, including specific referencing and the use of shared or recycled terms, aid meaning alignment. Similarly, imprecise or 'vague' language use was considered to undermine alignment. For example, discussing the following feedback extract an examiner suggested that the use of vague terms (i.e. 'pretty') left space for more clarification.

```
The standardisation marking is pretty accurate.
[E25 M1 26 5 1820]
```

Yeah, I guess that word 'pretty'. Yeah, 'pretty accurate' is quite vague, and perhaps could have done with perhaps a slightly more specific comment.

[Y2 E25 SR interview 32]

A specific concern voiced by examiners was that incorrect referencing in feedback messages led to confusion. In some cases, this included simple TL articulation 'errors' such as including the wrong script index number for a performance. For example, one examiner expressed this in their interview:

I'll open up [script] 2443, 1a. Oh no, that's not okay, well maybe it's the other question, or it's not there, right, which one is it? And you know ultimately it's just a slip of the finger but it then means you probably spent maybe five minutes or so

going through your other papers, and by that point you are just thinking 'which is the one?', and so, in that instance, that missing a number, probably, and you are already feeling a bit frustrated and stuff at that point.

[Y1 E18 SR interview 104]

6.1.5 Politeness and Face Management

Analysis also suggested that TLs' ambitions to use clear, unambiguous language sometimes created a tension with their desire to create greater social cohesion through language. According to Murphy & Rodriguez (2008) transactional distance is the psychological and communicative space of potential misunderstandings between teacher and learner inputs. My analyses suggest that TLs sometimes modified their language, using hedging devices, to reduce the transactional distance with the examiners. According to Lakoff, G. (1973) and Lakoff, R. (1973), hedging expresses tentativeness and avoids strong statements. At the same time, some argue that it dilutes the persuasive power of language (e.g. Huffaker, 2010; Johlke & Duhan, 2000; Mengis & Eppler, 2008). This tension was expressed by a TL who reflected on their own feedback message below:

```
4a I think this loses the 4th sig figs mark and no minus
sign.
[E5 M4 27.6.14]
```

I know bloody well it does [lose the 4th sig figs mark]. Rather than say 'It is' I say 'I think' because it is then inviting [the examiner]. It's not confrontational. And it also opens it if they come back to me and say 'No, I disagree'. Then we can have a normal professional discussion, which is what [this examiner] would do, which is fine.

[Y1 TL1 [E2E5] SR interview 443]

The tone that was cued by informal language use also helped to imply that reciprocity and dialogue are features of TL-Examiner relationships. This accords with findings reported by Harms & Roebuck (2010) and Hyatt (2005) and is demonstrated in the excerpts below:

```
Morning Examiner name, thank you for your first seed.
[E1 25.6.1048]
```

I think also it allows [the examiner] to feel involved. It's none of this 'Dear … Yours sincerely', or whatever. [To] build up a rapport.

[Y2 TL1 [E19E21E23] SR interview 350-364]

Yeah, definitely. I do like that … that sort of informality I guess, yeah. Yes, it's positive in a sense I think. Not so 'This is what you've done wrong'.

[Y2 E21 SR interview 39]

> This is out by quite a bit again but I can see the calls you have made (except for one mistake).
> [E8 6.9.1816]

I use parenthesis rather than a comma to change the language, it has become more informal. It's like the brackets and the formality is sort of, I guess there is thought in it here, saying, I feel like if I left that comment without the brackets and without dropping my tone a bit, it would sound as if I am saying, 'You are not giving this mark,' whereas 'actually I relate to you, I can see why you didn't'.

[Y1 TL2 [E8E11E12] SR interview 106]

> Hi *Examiner name*, I can give you feedback by Scoris messages.
> [E24 521.1653]

[The use of 'Hi'] certainly feels like [the TL] is establishing a relationship with me that is more of a support and a friendly relationship rather than a critical relationship.

[Y2 E24 SR interview 79]

Incorporating politeness into communication is another strategy that can be used to build social cohesion between participants (Brown & Levinson, 1987). In this way, strategies to promote politeness and reduce transactional distance share a common intent. Despite this, there appear to be few links in the literature between transactional distance and politeness studies. The work on transactional distance by Moore (1993) comes close to establishing a

link with politeness theories in the way that it suggests that distance can be expected to influence learner motivation, and in suggesting that this may also be influenced by personal learner preference and level of proficiency.

The feedback extract above, [E8 6.9.1816], demonstrates how the TL attempts to reduce the transactional distance with the less superior examiner through emphasising an equality of voice. This avoidance of 'bluntness' (which also has parallels with the tensions expressed around TLs' use of overemphasis, considered below) runs counter to research which suggests that bluntness contributes to feedback effectiveness (Bearman et al., 2013).

According to Bowe & Martin (2007), the expression of uncertainty is one indicator of symmetrical relations, and hedging adds to uncertainty (Salager-Meyer, 1994). This downgrading of status differences may discriminate examiner feedback from other professional contexts where superior and subordinate discourse has been studied. For example, previous research shows that politeness strategies are mainly used by subordinates when interacting with superiors in hierarchic workplaces (Morand, 1996, 2000). Moreover, when superiors use politeness they tend to employ positive politeness tactics, particularly with new and unfamiliar subordinates (Holmes, 2001; Holmes & Stubbe, 2003). It appears that examiner feedback runs counter to this observation, with TLs commonly displaying negative politeness in their interactions with examiners. It may be that this observation relates to the observation that examiner feedback supports a particular form of 'articulation work' (Johnson, 2015). Drawing on Strauss' (1985) concept, it is argued that TLs are very conscious of the need to ensure that competent examiners (or those who the TL believes will become so) are motivated to complete their marking tasks. The consequence of losing competent examiners from the workforce before marking is completed represents additional workload for the TL, and a concomitant strain on their relations with other examiners who then share the additional workload.

The use of emphasis in communication was another tension that was evident in the analysis. Some literature suggests that effective feedback involves the use of specific language that focuses on issues that matter most (Mengis & Eppler, 2008). According to this definition, highlighting or emphasising content would qualify as an effective feedback strategy, as it has qualities that are similar to bluntness. The feedback extract below demonstrates how a TL uses exclamation marks to reinforce a point of importance:

> 1b Most of this is waffle. At the bottom where you have put L2 is not creditworthy. It is just makes comments on creating concrete revetments and levee's to tackle the problem and give away fields to be used as floodplains SO WHAT !!!!!!!!!!!! There is no explanation of any kind here!
> [E15 12.6. 0911]

Analyses show that some examiners in my study felt that overemphasis could reduce message receptivity as it reinforced a sense of insecurity:

> *[In my past experience, the use of exclamation marks implies that] you are being shouted at in an email. You can feel it. As a marker you get your feedback message and you can feel quite, particularly early on in the thing as you are building up confidence a bit, you feel quite a bit apprehensive.*
>
> *[Y1 E5 SR interview 37]*

6.1.6 Collocated Interaction

Analyses broadly supported the notion that information conveyed through media rich communication channels (e.g. Trevino, Lengel, & Daft, 1987) positively supports common ground building. For the examiners in the study, rich communication possessed a variety of media qualities. Communication that contained a record of the interaction, or social or visual cues was considered to be useful for supporting participant alignment. The affordance of spoken feedback to cover a large amount of content in a relatively quick time was also mentioned by some examiners as a benefit of the mode. One examiner expressed this in the following way:

> *There is no voice to this person, so it's just their comments and sometimes you feel a bit exasperated by them. On the phone, it's immediate, the person is very friendly. It's much better like that, and I actually think it probably took less time, for [the TL] to do that.*
>
> *[Y1 E18 SR interview 77]*

Similarly, a lack of richness was considered to undermine alignment. A lack of richness describes the situation where participants lack the tools or language to convey a perspective (e.g. having access to only a restricted palette of digital annotations). This sentiment is

expressed in the examiner reflection on how the email mode limits their freedom of expression:

> *Sometimes I want to write something and I feel as I can't, I have to do it with an irrelevant ['IRRL' annotation] or a question mark ['?' annotation].*
>
> *[Y2 E29 SR interview 163]*

Media richness takes into account the way that information flows between participants and two-way interaction is a feature of more rich media[17]. The opportunity to have forward and backward information flow affords the opportunity to use questioning strategies. It has been noted that the use of questions in feedback can prompt critical reflection, (Hudson, 2014), and that the use of non-direct questions in particular can engage learners in their own learning through encouraging active listening (Ahmed et al., 2012). TLs employed questioning strategies to encourage a two-way dialogue so that points of view could be clarified through additional dialogue if necessary:

```
I hope that this helps?
[E18 M1 3.6.14]
```

> *So if he doesn't understand then I really would like him to get back to me. I think he would actually, if he is stuck.*
>
> *[Y1 TL3 [E16E18] SR interview 84]*

6.1.7 The Recipient Having Trust in the Feedback Giver

The dominant theme that was theorised from the literature was that the existence of trust between interacting participants would influence the reception of feedback communication and common ground building. The evidence suggests that the TLs were using their feedback messages to construct a relationship of trust with examiners. The extract below demonstrates how a TL uses reference to the personal pronoun 'we' to emphasise the collegiate nature of their relationship with the examiner:

[17] Although I need to point out at this stage that this research is not a study into the affordances of information technology which has been dealt with elsewhere (e.g. Bower, 2008; Conole & Dyke, 2004), instead my research is interested in looking at how communication takes place within these established affordances.

```
…we were generous in a and gave the deltaH mark.
[E2 M4 28.6.14]
```

Because it's not a personalized mark scheme where I did it, it's the team. So it was 'we gave it' so 'we are wrong.' It's not a me against them, it's the team that is the hierarchy, if you wish to use that expression, gave that mark.

[Y1 TL1 [E2E5] SR interview 222]

Trust appeared to relate to the interaction environment and, more specifically, to the quality of the communication that was possible between participants. Where feedback lacked rationalisation, it was possible for examiners to perceive a basis of distrust underlying their relationship with the TL:

I've worked with other team leaders in other situations who have been really poor at giving feedback. I mean the worst thing is when someone rings you up and says, 'Yeah, it's fine, it's fine', because you think well 'No, no, it can't be', so in a way that makes me think 'You haven't even looked at it'.

[Y2 E26 SR interview 128]

This point links with the findings of Morgan & Symon (2002) who note that open workplace communication leads to trust. A sentiment within the examiner extract above also alludes to the sense that feedback can contribute to perceptions of professional isolation. Huffaker (2010) argues that effective leaders in remote work reach out to others through their communication, and engage in relationship development. In the following extract, an examiner reflects on how a TL had managed to deal with this in their feedback communication:

And more recently there's things that have come up and [the TL] had said comments which have obviously shown that she has similar issues to me in terms of some of the things that have come up.

[Y2 E29 SR interview 155]

Sharing personal perspectives can be an important part of this engagement, and it has been noted that this can lead to social closeness and the reduced potential for professional attrition (Owens et al., 2009).

As a counter to professional isolation, it is possible that examiners draw on additional resources when unsure about acceptable mark scheme interpretation. As one examiner stated:

> *I had originally assumed that that was the right way to mark it and a colleague persuaded me that I was wrong. So I then marked it in the way in which my colleague marks it, and he's done a lot more of these than I have. And then [the TL] came back to me and said, 'Actually, you've got that wrong' and I thought, right, I was right to start with*
>
> *[Y2 E24 SR interview 134]*

This point accords with the work of others who have reported how novices who are insecure in their professional learning gain confidence through accessing information from trusted peers (Bearman et al., 2013; Burke et al., 2009; Segoe, 2013). This has a consequence for hierarchic information systems. Trust was also considered to be compromised where multiple sources of information (e.g. 'parallelism', Münzer & Holmer, 2009), or mixed messages, existed in a line of communication. This is discussed by an examiner whilst reflecting on a feedback extract:

```
It's a messy answer but on reflection I agree we should
have awarded an analysis mark for supply meeting demand.
[E24 M1 21.5.1806]
```

> *[The TL] told me on the telephone that when we've got that phrase, 'increasing supply meets increasing demand' then, because it's good as far as the Exam Board's concerned, you can award three marks because of the 'increasing supply' the 'increasing demand' and the word 'meets' that joins the two phrases together in the middle. But, they then give me a standardisation script and I mark according to what [the TL] told me on the phone and they've ignored it. So that's why I found it very difficult then to apply some of the other principles that they were asking me to apply.*
>
> *[Y2 E24 SR interview 55]*

The examiner explains that this feedback message embodies a conflict where the examiner is left uneasy about what information to trust *at a particular given time*. This is problematic as it undermines the examiner's confidence in their on-going decision-making, and coheres with the work of Bosley & Young (2006) who report that single stream (i.e. 1:1) dialogues between experts and learners are most effective for learning.

Power differentials between the feedback source and the recipient also had a reported influence on information sharing. Higgins, Hartley, & Skelton (2001) report how feedback that involves a mixture of power and judgement-making has a heightened emotional impact. In this situation, a subordinate may feel reluctant to share information as it risks them appearing incompetent (Bartolomé, 1989). There is evidence in my data that hierarchical differences in the professional status of the participants can act as a limiter on their relationship and trust building, as there is social pressure that hinders free communication. For example, an examiner reflects on the embarrassment that they feel when they present work that is below standard:

```
1ai I am not sure what you are doing here as you have 5
ticks for a start? Only give one tick per mark and look
at the mark scheme to see how they are allocated.
[E18 M1 3.6.14]
```

> *You feel embarrassed, I suppose. You put five ticks because you are trying to make it perfect and you think, right she has told me before to only put ticks in the right place. I have done something wrong here, to the most simple question, that's the kind of thing that if one of my students did, I would call them up, and think they were a bit of a dumbo.*
>
> *[Y1 E18 SR interview 43]*

6.1.8 Positive Information Feedback

Analyses suggest that the inclusion of positive information in feedback messages is perceived as being motivational and more prone to being well-received. Similarly, negative content had an adverse emotional impact on examiners and left them feeling anxious:

```
If you could please review the comments I've made on
each script, adjust as appropriate and submit that would
be great.
[E26 M1 21.5.1759]
```

> *Some of the feedback on standardisation scripts has been a bit more blunt. But then that can feel a bit of a jolt because, you know, things are going 'Hi, that's a great script, well done', 'that script is very accurate', and then you suddenly get a message that says 'This one was out of tolerance, have a look at this again'. Maybe I'm just being sensitive but you kind of think 'Oh no, I'm going wrong'.*
>
> *[Y2 E26 SR interview 118]*

Some TLs invoked empathy in their communication through the use of positive politeness strategies. This feature is noted in other feedback studies (e.g. Sussman & Sproull, 1999), and it is suggested that empathy may help to reduce the transactional distance that is generated when negative information is communicated (Ackerman & Gross, 2010). This feature of examiner feedback has been reported elsewhere (Johnson & Black, 2012a), where it was tentatively proposed that a feature of effective TL practice is the foregrounding of a view of the examiner as a learner. A corollary of this is that professional face implicates emotions that can be undermined by professional criticism. This is reflected in the examiner interview extracts below:

> *It's quite a difficult thing being judged by your peers like this, because as somebody who has been teaching the subject for a long time, suddenly somebody is telling you that maybe what you view as being the way it is isn't maybe quite right. It's quite an emotional thing to receive that feedback.*
>
> *[Y1 E11 SR interview 43]*
>
> *Well you're basically being told that you're doing it wrong or you've done that bit wrong, but as a "professional", in inverted commas, it's very difficult to take that.*
>
> *[Y2 E26 SR interview 122]*

In line with the work of Mathisen (2012), it appears that feedback that prioritises the importance of positive information helps to induct professionals into a community. This is reflected in one TL's comment:

> *[I don't pass on all of the errors because] you've also got to bear in mind their professionalism… his marking was still excellent so therefore that kept him saying 'Yes, I'm still a professional and I know what I'm doing.'*
>
> *[Y1 TL1 [E2E5] SR interview 78-93]*

The TLs in my study dealt mainly in negative information (see results in Chapter 5), but they often considered how best to balance their communication of positive and negative information. This is a feature of desirable feedback noted by Tang & Harrison (2011). The positioning of positive information was also perceived as having a potential influence on motivation. Foregrounding positive information was perceived as ensuring that the recipient would be receptive to the message. This corresponds with the notion of the feedback sandwich reported in Archer (2010). The feedback extract below show a typical opening used by TLs, with their rationale for this action:

```
Thank you for your standardisation scripts which were
excellent.
[E23 M1 19.6.1045]
```

> *In this case they were excellent, and so I told her so, rather at the beginning. So therefore, she wasn't on the back-foot, shall we say, when she was looking at it.*
>
> *[Y2 TL1 [E19E21E23] SR Interview 82]*

6.2 Attaining Alignment

In the previous section, I considered the issue of convergence and divergence from a theoretically grounded perspective. This helped to link the feedback data in my study to the broader literature and suggested that there was a link between feedback and trouble sources. In this next section, I want to use the direct evidence from examiners' interactions to consider how they overcome divergence around particular trouble sources. To do this I look at how the participants attain resolution through feedback discourse. Resolution is a public indication that a repair to a trouble source has taken place, and this constitutes an element of *Impact*

(which is one of the areas of interest for my analysis alongside *Content*, *Time*, and *Joint Intellectual Action*).

My analysis in this section includes two elements. In the first instance, I use participants' indications of resolution across different interactional cases to identify common patterns across these cases (section 6.2.1). For this analysis, I draw on the coded evidence of resolution (see Chapter 5.2.1).

A second element of analysis is necessary because of the problem of tacit agreement (i.e. where there may be no tangible trace of resolution). According to a legal definition, a tacit agreement is implied or inferred by a participant where another participant refrains from contradiction or objection (Garner, 2014). The tacit notion is also an important element in the ethnomethodological perspective (Lynch, 2001) and underpins Conversation Analysis, thereby informing my study. According to this perspective, the unfolding nature of discourse leads participants to establish working hypotheses about the state of other participants' understanding at a particular time. These working hypotheses are confirmed or refuted through utterances as discourse develops, and this information forms the content of shared common ground. As a result, established discourse (that builds on the foundation of earlier discourse) contains latent (or redundant) content that does not require explicit re-statement by participants.

In my study context, there are situational features that accentuate the pressure for examiners to move quickly towards maximising redundancy in their discourse. For example, a number of workload and focus issues conspire to ensure that TLs do not generally use discourse to communicate agreement. These issues include:

- Tight marking task completion deadlines for both participants, a pressure on feedback giving that is noted in other contexts (e.g. Chetwynd & Dobbyn, 2011) (workload);
- The effort required to open up the electronic messaging system and compose a message (workload);
- The principal discourse focus being on remediating marking error (focus);
- Concern that additional interaction around agreement may detract from the effect of discourse that deals with disagreement (focus).

A consequence of such redundancy is that if feedback is generally negative in character (i.e. picking up areas of examiner error), a lack of feedback can be used by examiners as a working hypothesis to indicate that they are marking acceptably.

Latent content is a challenge for analysis as it is only through the participants' implied allusion to such content that its existence can be evidenced. Resolution can occur without it being demonstrated as a publically accountable statement (e.g. the participants in discourse do not formally acknowledge their shared agreement). This means that analysis inevitably relies on the use of interpretation, drawing on the historical development of the discourse, and ideas represented within it.

This leads to the second approach that I use to explore resolution building. In section 6.2.2 I explore the case histories of problematic items where a TL gives repeated feedback over the course of the marking period. Countering concerns that resolution might exist but without any observable evidence, this approach does not rely on publically stated resolution and focuses instead on the existence of interaction as being a de facto indicator of disagreement. The assumption underlying this approach is that the maintained existence of feedback is an indicator that differences in understanding remain unresolved (and a trouble source persists). This builds on the premise that examiner feedback discourse tends to focus on those areas where disagreement exists.

6.2.1 Resolution Analysis

In this section I present analyses of a sample of feedback data where the participants interact around a problematic item.

As I outlined earlier (Chapter 4.3.2), these analyses are largely qualitative in nature and do not rely on indicators of statistical significance. This is because the study of subsamples using qualitative analysis can reveal things that would not otherwise be recognised due to a 'lack of statistical power' in the context of a conventional quantitative analysis (LeCompte & Preissle, 1993), and a 'unique-case selection' allows close analysis to highlight the processes that operate within a select group (and is therefore not based on statistical relations).

6.2.1.1 Content, Time, Joint Intellectual Action, and Impact (Resolution)

For this analysis, I identified 329 instances of feedback data where there was evidence of two-way interaction. From these, 55 episodes were sampled where coded evidence of resolution from initially diverse perspectives was recognised (see Chapter 5.2.1). It is worth

re-stating here that resolution is the key indicator that the features of effective feedback, outlined above, have, in fact, led to a positive outcome. In the context of examiner feedback, resolution is a signifier that alignment has been attained by the examiners.

To gain an overview of the patterns of discourse around resolution, the next level of analysis identified six discourse codes from two code families that co-occurred with the resolution code. These were *Rationale for credit*, *Location of credit*, *Request information*, *Examiner rationale* (all from the 'GI' [Giving Information] theme), and *Historic reference* and *External reference* (from the 'Br' [Bridging] theme). According to my theoretical framework, Giving Information codes relate to feedback *Content* and Bridging codes relate to *Joint Intellectual Action*.

A visualisation of the interactions between the codes, generated by the MAXqda textual analysis software (VERBI Software – Consult – Sozialforschung GmbH, 2013), is presented in Figure 6.1. For the purposes of discussion, the visualisation also includes information about a code that serves as a point of contrast with the other codes (*Mark statement* [GI]). In the visualisation, the most dominant links between codes are indicated by the weight of the connecting line. It should be noted that the visualisation does not represent causal interactions, but rather the relations between codes that culminate in resolution acts. This means that, for example, interaction episodes may be instigated by the request for information (from a TL or examiner) that leads to rational sharing, or vice versa.

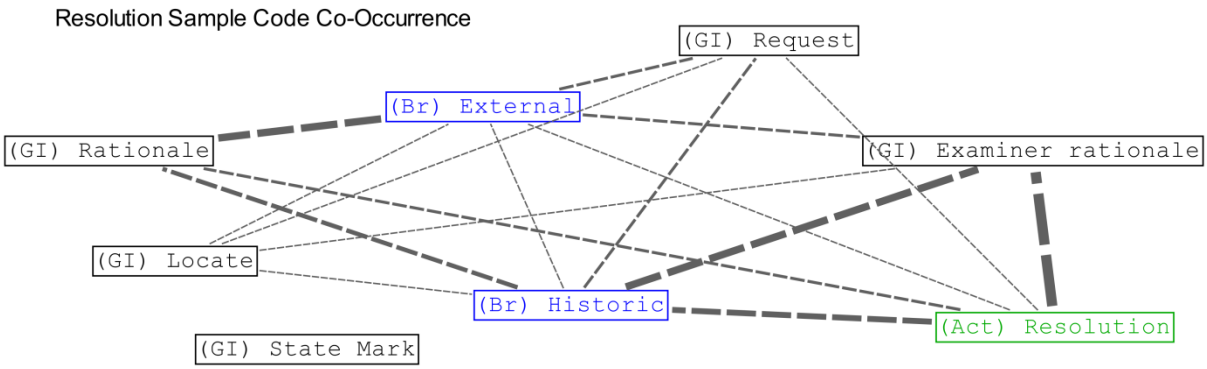

Figure 6.1: Codes that Interact Around Resolution

The visualisation is useful for highlighting a number of issues. Firstly, it is noteworthy that sharing a rationale for credit, or sharing the location of credit, has more impact than stating a mark. Secondly, the link between rationale sharing and resolution interconnects with external and historic bridging work (bringing together information). Finally, there is a connection

between the examiner sharing their rationale and resolution. To gain an overview of the chronological patterns of discourse around resolution the next level of analysis identified the codes that co-occurred or immediately preceded resolution.

The visualisation of the resolution sample co-occurrence (presented above) made it possible for me to look at the feedback discourse in a new way, and to identify some patterns around the chronology of interconnecting codes. Interpretation of the visualisation suggests that there are four patterns that describe how Giving information (*Content*) interconnects with Bridging (*Joint Intellectual Action*) to attain Resolution (*Impact*). Each of these patterns is presented in Figure 6.2.

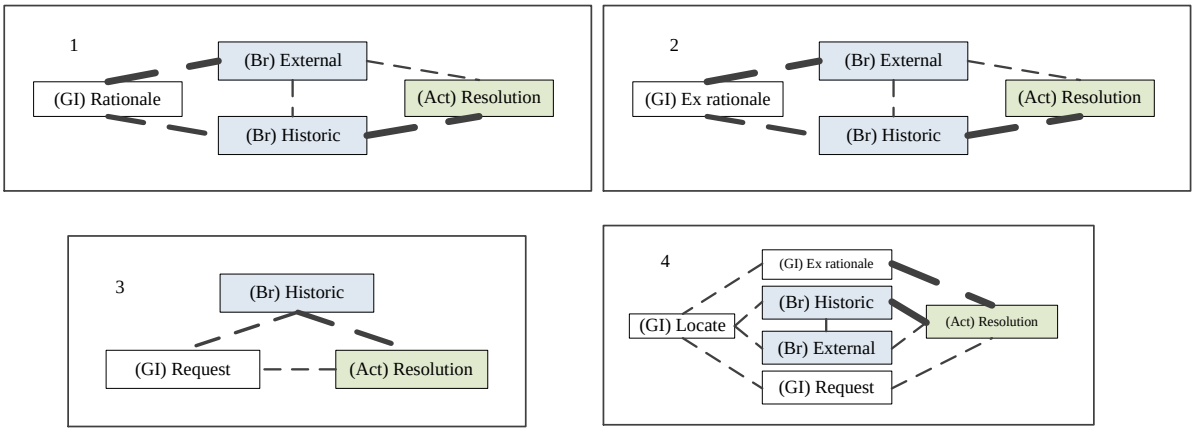

Figure 6.2: Resolution Patterns

To exemplify each of these patterns I present extracts of feedback discourse. The presented data seek to illustrate the key points of the analysis but have, once again, been chosen judiciously because of the constraints of the thesis word count. For reasons of clarity, I present the first of these extracts in two formats, as text only (Figure 6.3) and as an annotated extract that includes the coded moves that resulted from my analysis (Figure 6.4). Moves are deliberate actions that aim to influence discourse and their analysis has been used elsewhere as a focus for studying learning interactions (e.g. Krussel, Edwards, & Springer, 2004; Lefstein, Snell, & Israeli, 2015). As I outlined earlier (Chapter 4.1.2/4.2.2), thematic coding allowed double coding of text segments to account for the multiple purposes that a singular text segment might be achieving (e.g. a highlighted word could be interpreted as both accentuating an important concept as well as conveying something about the authority that is implicit in the relationship that the participants are constructing and maintaining through their

use of the word). In the example below the moves are identified by a number (1-5). For the later extracts, I present the feedback data as text only to aid reading ease of reading.

6.2.1.1.1 Resolution Pattern 1: TL Rationale>Bridge>Resolution (Extract 1):

In the first extract (Figure 6.3) a TL and an examiner are discussing an Economics question.

From: TL: 04/06/2014 14:10
Hi *Examiner name*,
Can you please review this, particularly 6b? It's quite a bit out of tolerance.
5aii - says the models are inelastic, which doesn't get a mark as it needs to be demand.

>From: EXAMINER: 04/06/2014 2:55:14 PM
Hello
Q5aii) I have marked it based advice below i.e the figure/value is income inelastic. Is this not acceptable?
The joys of not having a meeting.
Thanks

>>From: TL: 04/06/2014 17:18
Re:RE: Standardisation 1 Script 3
I can't find the script you are referring to for Q5aii - can you type in whatever the candidate has written for it so I can explain why it didn't get a mark?
Thanks,
TL name

>>>From: EXAMINER: 04/06/2014 8:26:07 PM
'As the income elasticity of demandis 5 it means that its a normal good hence as incomes fall its deman will fall and also as it is greater than one it is income elastic hence its more of a luxury. Hence during recession Airfix must be awareaware that its demand will fall'.
'As it is ... ' assume this is not enough to refer to income elastic and YED figure.
Is it different if just say the vaue figure ..is income inelastic - this would be more the issue than the answer above as this is less linked to the figure itself.
Thanks

>>>>From: TL 05/06/2014 12:47
Re:RE: Re:RE: Standardisation 1 Script 3
The problem is that they aren't clearly saying the value/figure/YED is inelastic, which would be fine, but have got some analysis in between 'income elasticity of demand' and 'it is income elastic', meaning it is not clear whether 'it' refers to the value or the good, especially given this follows them saying 'it is a normal good', implying they are treating 'it' as the good itself.
Hope that makes sense!

[E8 4.6 1410/E8 4.6 1718/E8 5.6 1247]

Figure 6.3: Resolution Pattern 1: TL Rationale>Bridge>Resolution (Extract 1) [Text Only]

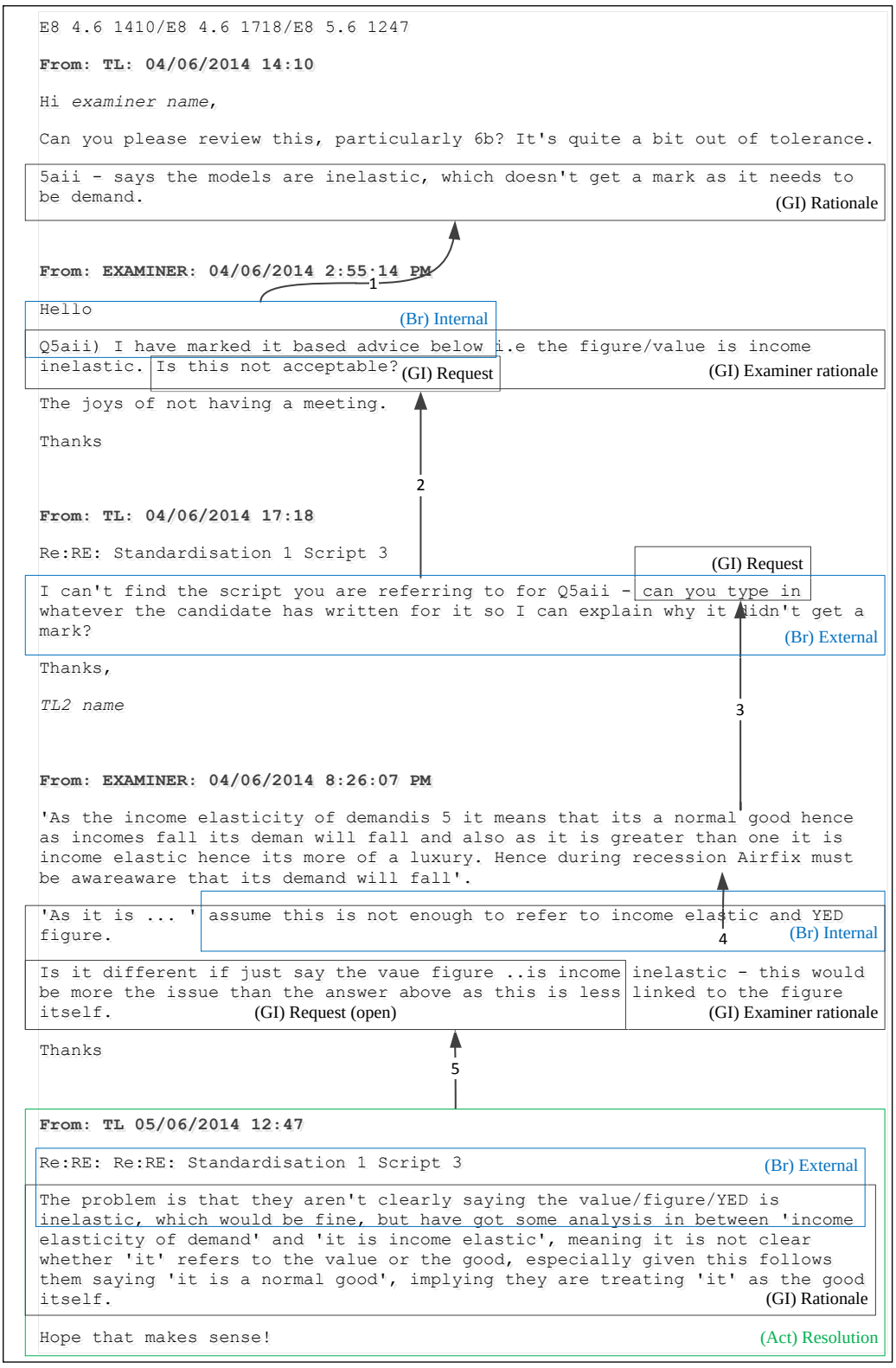

Figure 6.4: Resolution Pattern 1: TL Rationale>Bridge>Resolution (Extract 1) [Annotated Text]

In the Economics question being discussed in this example the problem centres on a common candidate error. The question asks candidates to explain what a specific ‘income elasticity of

demand value' for a model kit indicates to a company (Airfix). The mark scheme outlines the marking rationale for the question (Figure 6.5):

> A candidate should explain what the figure shows for three marks such as:
>
> the YED is INELASTIC (1)
> the product is, therefore, insensitive to changes in income (1) and so is relatively "recession-proof" (1)
> a NORMAL good (1)
> as the economy starts to grow and incomes rise, the quantity demanded will rise (1) by a smaller percentage. (1)

Figure 6.5: Economics Mark Scheme (1)

In many cases, it is found that candidates incorrectly reference the 'model' or 'the good' as being income inelastic, whereas the correct response should refer to the 'demand' or 'the value' as being income inelastic.

There are five moves in the feedback sequence. The sequence starts with the initial feedback message and involves an iterative refinement of information to culminate in a point of resolution that locates a problematic pronoun in the candidate's response. In the first move the TL articulates a rationale for credit, using the phrase 'as it needs to be demand'. The examiner uses an internal bridge - a reference to previous information - to explain their thought position. This leads to the second move, which is prompted by the examiner's use of a question (`Is this not acceptable?`) and signals that the rationale is not adequate and that misalignment of perspectives still exists. The positioning of the question following a stated thought position places the onus on the TL to respond with additional information.

In the second move the TL requests additional information (`can you type in whatever the candidate has written`) that can then be used as a resource for shared reflection. This request is an external bridge as the participants begin to draw on an external reference document (the shared examination script).

The third move is a direct response to the request for information as the examiner presents the candidate's wording for reflection. This information is then the basis for the fourth move where the examiner takes the opportunity to use the candidate's text to articulate their point

of view to the TL. Here, the examiner is using an internal bridge to draw the TL's attention to important features of their thinking via the text. This is then reinforced by a conjectural point that is raised in the form of a question. This leads to the fifth move, as the TL addresses this question directly using an external bridge to reference the candidate's text. This move leads to the identification of a specific pronoun ('it') as the cause of the misaligned perspectives that related to the candidate response.

6.2.1.1.2 Resolution Pattern 1: TL Rationale>Bridge>Resolution (Extract 2):

In the second extract (Figure 6.6) a TL and an examiner are discussing a different Economics question.

From: TL 25/05/2014 16:37

Hi *Examiner name*,

7 - we kept this answer in L2. Only L1 should be awarded for a market failure diagram. We didn't give the first L3 as that is not saying a TAX causes the private cost curve to shift. We didn't give the second L3 because the candidate is saying a tax is given to producers, which isn't right.

Hope this all makes sense. I'll send this script back to you for you to have a look at and make changes before resubmitting. Let me know if anything is unclear.

Thanks,

TL name

>From: EXAMINER: 25/05/2014 4:52:17 PM

hi,

Can i just ask about question 7- are you saying that entire essay was kept to level 2 as there is discussion ie pros and cons/conparison?

Examiner name

>>From TL: 25/05/2014 17:19

Re:Standardisation Feedback

Yes. A discussion of pros and cons without clear economic analysis of a tax (L3B1) and then analysis of how this solves the market failure associated with the negative externalities of production (L3B2) gets capped at L2.

[E9 25.5 1637]

Figure 6.6: Resolution Pattern 1: TL Rationale>Bridge>Resolution (Extract 2)

In this extract, the TL and examiner discuss an Economics essay response in a practice marking script. The 18-mark question asks candidates to discuss whether indirect taxation is

the most effective policy measure to correct the market failure arising from the negative externalities of production. The question implies the need to consider the case for and against direct tax use, and a key discriminator for a performance to move from Level 2 [L2] to Level 3 [L3] (and being awarded more than eight marks) is the inclusion of economic analysis. This is articulated in a section of the question mark scheme (Figure 6.7):

> Level 2 responses may offer an unexplained diagram or raise, but not analyse, some of the following:
>
> - indirect tax will increase supply
> - imposing an indirect tax will raise the price of the products.
>
> An answer cannot progress from L3 B1 to L3 B2 without analysing how a tax changes the QUANTITY produced.

Figure 6.7: Economics Mark Scheme (2)

There are three moves in this extract, with the TL steering the examiner to focus on the concept of 'economic analysis' as a key indicator of a L3 performance.

In the first move, the TL outlines their rationale for why the answer is L2, with this centring on the use of the word 'because' and an appeal, using an external bridge, for the examiner to reflect on the shared script. The TL finishes the message with an invitation for the examiner to respond if there is anything unclear to them. This request forms the basis of the second move, which takes the form of an examiner request for clarification.

In the second move, the examiner expresses their perspective in the form of a question. In the body of this question they use external bridge when they make a reference to the mark scheme (`i.e. pros and cons comparison?`). The third move sees the TL using this reference to the mark scheme as a common resource which they use to clarify the particular requirement to involve '`clear economic analysis`' as a discriminator for L3 performance. This requirement is not actually stated explicitly in the mark scheme.

6.2.1.1.3 Resolution Pattern 2: Examiner Rationale>Bridge>Resolution (Extract 1):

In this extract (Figure 6.8) the TL and examiner discuss a 25-mark Geography question where candidates need to explain how climate influences the physical landscape of cold environments with reference to one or more located examples. The examiner is unsure about how to interpret the knowledge and understanding assessment objective (AO1) of the mark scheme.

From: Examiner: 28/07/2015 08:33:24 AM

Dear *TL name*

Bit of a quandry as I want to give this final essay in the middle level 2 for AO1 as it refers to climate effects but think it may deserve middle level 1 as it does not talk about landforms.

Help please

Examiner name

>From: TL: 28/07/2015 09:43

Hi again *Examiner name*

This script is very weak and should not get into L2 your second opinion seems about right. See clip from markscheme below for L1:-

'Limited knowledge and understanding of the processes influencing landscape features.

Cause-effect links are limited or absent. There is limited exemplification.'

Happy marking

TL name

>>From: EXAMINER: 28/07/2014 10:00:36 AM

Thanks

[E28_728.0833_Q4]

Figure 6.8: Resolution Pattern 2: Examiner Rationale>Bridge>Resolution (Extract 1)

There are three moves in this extract, with the TL steering the examiner to focus on one element of the mark scheme to help them to resolve their insecurity about the quality of the assessed performance.

In the first move, the examiner outlines their rationale for why they could perceive the candidate performance as being either Level 1 or Level 2, centring on the use of the word 'as'. In this segment, the examiner uses an external bridge when they refer to an element of

the performance (`it refers to climate effects`). This leads to an appeal for help, which is the beginning of the second move.

In the second move the TL steers the examiner's perspective through the use of an external bridge where they make a direct reference to the mark scheme (`Limited knowledge and understanding…`). This leads the examiner to indicate resolution in the third move (`Thanks`).

6.2.1.1.4 Resolution Pattern 2: Examiner Rationale>Bridge>Resolution (Extract 2)

In the next extract (Figure 6.9) the TL and examiner use the telephone to discuss the income elasticity of demand question that was discussed in Resolution Pattern 1/Extract 1. Again, the focus is on ensuring that candidates correctly reference the location of income inelasticity.

```
001   TL    ::: yes, so ::: with 5aii, the air fix model is income
002         inelastic, we're being very strict on this. It has to [be
003   EX                                                      [I was
004         about to] ::: sometimes they say it's fine if you say
005         inelastic but then I notice on the mark scheme that it's
006         like you can't say that at all
007   TL    Yes, you have to say demand is income inelastic. If it's
008         demand is inelastic that's not enough. The only way they
009         can get around it is if they talk about the value or the
010         figures or YED being inelastic,that's ok because that's
011         like a correct sentence, [but]
012   EX                             [ok]
013   TL    But a plane can't be inelastic
014   EX    So you gave it 2 out of 3 then
015   TL    Yeah.

[E9 25.5 T]
```

Figure 6.9: Resolution Pattern 2: Examiner Rationale>Bridge>Resolution (Extract 2)

There are four moves in this extract, with much of the discussion using external bridging to focus on the refinement of already agreed thought positions. In the first move, the TL introduces a question for which the examiner and the TL had reached a different mark outcome. Within this move the examiner interjects to outline to the TL that they were aware of where the discrepancy might be. To do this, they use the external bridge of referring to the mark scheme (`then I notice on the mark scheme…`). In the second move, the TL

builds on this reference, expressing agreement with the examiner. This move then becomes the basis of the third move, where the TL reiterates the reasoning of the mark scheme. The fourth move involves the final resolution of the issue, with the examiner checking that their understanding of the TL's perspective – expressed in the form of marks awarded to the performance – is correct.

6.2.1.1.5 Resolution Pattern 3: Request>Bridge>Resolution (Extract 1):

In this extract (Figure 6.10) the TL and examiner discuss a 6-mark Chemistry question where candidates need to describe the oxidation reactions of propan-1-ol when using a suitable oxidising agent and include reagents, observations and equations in the answer. The mark scheme includes guidance that states: 'Observations: Orange to Green'. In this feedback episode, the TL refines the examiner's understanding of an unwritten nuance of the mark scheme.

From: TL 19/06/2014 21:11

Hi *Examiner name*

Many thanks for the first batch of 5 script for reviewing. A couple were outside of the tolerance of +/- 5, and will be sent back to you for amendment and resubmission, please.

The following comments are to help clarify some of the points on the marking scheme:

ID: 648231789: 7d we gave the colour change - altho' I can understand why you didn't.

ID: 649878483: 3a Both marks score here - both lines start from the reactants - baseline.

…[15 ADDITIONAL FEEDBACK LINES REMOVED]

ID: 649108490: 1a(i) we gave the 2nd mark here - 'differs by CH2 each time'.

ID: 647304860: 7d The aldehyde equation isn't balanced.

Hope this helpful - look forward to the resubmitted scripts and the next five, please.

Best regards

TL name

>From: EXAMINER: 20/06/2014 11:08:31

Hi *TL name*

Many thanks for helpful comments. I have one query - if I understand correctly, we gave the colour change in 7d for 'orange' in 789, but not for 'green' in 860. Is this not a bit of a contradiction?

Examiner name

>>From: TL 20/06/2014 16:58

Hi *Examiner name*

We did give the colour change in 789 for orange to green (altho' the answer states the carboxylic acid went from orange to green) - probably what is seen in the reaction vessel.

But in 860 is was just 'green' not mention of from what colour.

TL name

[E6 19.6 2111]

Figure 6.10: Resolution Pattern 3: Request>Bridge>Resolution (Extract 1)

There are two moves in this extract, with the TL steering the examiner to focus on one precise element of the mark scheme that requires candidates to include explicit reference to both colours (orange and green), although this is not overtly stated in the mark scheme. In the first move, the TL provides information on where credit should be located for Question 7d in a number of marked scripts (amongst many others).

In response to the feedback the examiner bridges across the information, forging a historical linkage between two elements of the TL's feedback message. This linkage provides a resource for the examiner to use to outline their perspective. This rationale sharing takes the form of a request (`query`) for additional explanation to overcome an apparent contradiction in the TL's feedback message. The examiner request leads to the second move, where the TL uses a reference to the performances mentioned in the original feedback to outline their rationale in more detail and to serve as a resolution to the issue.

6.2.1.1.6 Resolution Pattern 3: Request>Bridge>Resolution (Extract 2):

In the next extract (Figure 6.12) the TL and examiner discuss a 4-mark Economics question where candidates need to use a supply and demand diagram to show how consumer surplus would change if the supply of a product decreased. The mark scheme outlines how the marks can be allocated (Figure 6.11)

Up to four marks for a diagram which clearly illustrates a loss in consumer surplus and rise in price:

- Supply and demand curves with shift left of supply (1)
- New coordinates for higher price and lower quantity (1)
- area of original consumer surplus (1) e.g. A+B
- area of new consumer surplus (1) e.g. A
- shaded or labelled area of difference (1) e.g. B
- statement or label "lost" (1)

Figure 6.11: Economics Mark Scheme (3)

From: TL: 29/05/2015 17:08

3b - the original consumer surplus is correct so one mark for that.

>From: EXAMINER: 29/05/2015 5:20:19 PM

3b Zero marked owing to supply going in the wrong direction. Is this a version of the OFR that is being applied here, in other words candidate gets the mark for showing where the original consumer surplus is according to his/her application despite its being incorrect? To me this candidate got the answer completely wrong. Of course, I will apply this rule to further scripts, however.

>>From: TL: 29/05/2015 21:00

3b is an interesting one. Imagine if their diagram were exactly the same but the supply curve had been shifted the right way. Obviously this would change the rest of the answer, but the original consumer surplus area actually wouldn't change at all. So even though the diagram looks entirely wrong, the initial consumer surplus is actually correctly identified - it's what happens next that doesn't make sense because they've shifted the curve the wrong way! So no OFR mark - just a mark for correctly identifying the original consumer surplus from an initial price before any change takes place in the market.

>>>From: EXAMINER: 29/05/2015 9:05:52 PM

Absolutely! Thanks, *TL name*.

[E24_529.1714_Q3]

Figure 6.12: Resolution Pattern 3: Request>Bridge>Resolution (Extract 2)

There are three moves in this extract, with the TL building on the information that they had originally shared but which proves to be inadequate for the examiner's needs. In the first move, the examiner responds to the TL feedback by making a direct request for information using an external bridge reference to a mark scheme concept (`Is this a version of the OFR [Own Figure Rule]`). The second move is a response by the TL to the request. In this move the TL uses a conjectural phrase (`Imagine if…`) to share their rationale. To do this, they use the original candidate response as a resource for extending the concepts to a context that has the examiner has not yet experienced. This move serves to extend the context to becoming more generalizable by ensuring that the examiner focuses

exclusively on '`the consumer surplus area`'. The third move is used by the examiner to indicate agreement with the TL's stated position.

6.2.1.1.7 Resolution Pattern 4: Locate>Examiner Rationale/Bridge/Request>Resolution (Extract 1):

In this extract (Figure 6.14) the TL and examiner discuss a 9-mark Geography question where candidates need to explain how river basins can be protected from the effects of flooding. The mark scheme (Figure 6.13) includes guidance that states to reach the marking highest band (8-9 marks) the candidate…

> Uses clearly identified example(s) to explain how protection is achieved. Causal links between method and at least two effects are clearly explained. Answer is well structured with accurate grammar and spelling. Good use of appropriate geographical terminology.

For performances at the lower band (5-7 marks) the candidate…

> Gives clearly identified example(s) to explain how protection is achieved. Cause-effect links are stated but explanation may not be clearly linked to effects. Answer has sound structure but may have some errors in grammar and spelling.

Figure 6.13: Geography Mark Scheme (1)

In the following feedback episode, the TL moves from locating the credit in a candidate's performance to outlining their perspective, using external bridge referencing.

001 **TL** Right, question 1c. The middle bit, The Thames Estuary.
002 That's 9. So read through it.
003 **EX** Right? ::::::::: Yeah?
004 **TL** You've got it?
005 **EX** So that one section there is worth 9?
006 **TL** That's 9
007 **EX** Right
008 **TL** Yeah
009 **EX** Oooh! > Right.
010 **TL** Because if you can see, a lot mention the Thames Barrier
011 but haven't a clue that it's tidal surges that it's
012 against
013 **EX** Yes
014 **TL** They just say to regulate the water, something woolly
015 like that. But this one. He's got it spot on. *High tide*.
016 *Stormy condition. Risk of flooding*. And then he's
017 actually developed the risk to The Canary Wharf and all
018 that
019 **EX** The economic, yeah
020 **TL** So it's a really good answer so that is worth 9 on its
021 own.
022 **EX** Right, I'll add that to 9 now.

[E15 22.6 T]

Figure 6.14: Resolution Pattern 4: Locate>Examiner Rationale/Bridge/Request>Resolution (Extract 1)

There are three moves in this extract. In the first move, the TL states the location of credit within a shared performance script. The examiner repeats the TL's view in the form of a question, suggesting that they disagree with this evaluation. This question leads the TL to the second move, where they articulate their reasoning for the mark awarded. To do this, they directly reference the wording of the candidate's script as a form of external bridge (lines 015-016). Move three suggests that this articulation supplies the examiner with the adequate amount of detail as they not only re-voice a concept that was implicit in the TL feedback (the economic), but they also decide to amend the marks that they allocate to the script.

6.2.1.1.8 Resolution Pattern 4: Locate>Examiner Rationale/Bridge/Request>Resolution (Extract 2):

In the final extract (Figure 6.16), the TL and examiner discuss an Economics essay response. The 18-mark question asks candidates to discuss whether indirect taxation is the most effective policy measure to correct the market failure arising from the negative externalities of production. For a performance to move from Level 2 [L2] to Level 3 [L3] the candidate

needs to clear L2 ‘trigger points’. These points are articulated in a section of the question mark scheme (Figure 6.15):

Band 2 points:

- connecting shift leftwards of supply with correcting “negative externalities of production”
- corrected equilibrium reduces over-production thus solving market failure

Figure 6.15: Economics Mark Scheme (4)

In this feedback episode, the TL moves from locating the credit in a candidate’s performance to outlining a simple reference from the mark scheme, helping the examiner to understand how they should mark in the future.

From: TL: 08/06/2014 15:04

This is quite a long way out of tolerance, can you please review? The essay hits a clear Band 2 trigger in the mark scheme, which is why it is able to progress through to the top.
5b - only 1 mark as helps with setting prices/maximising revenue is an either or mark.
6b - reduced overproduction gets this answer through Band 2 and it goes onto get 18.

>From: EXAMINER: 08/06/2014 3:32:42 PM

Hi *TL name*
OK. I didn't think it was clear enough for Band 2, but I'll ensure that I reward it if I see 'reduced overproduction' again.
Thanks
Examiner name

>>From: TL: 08/06/2014 16:03

Yes it's an easy route in really, but very few candidates are putting it so those who do should be rewarded.

[E11 8.6 1504]

Figure 6.16: Resolution Pattern 4: Locate>Examiner Rationale/Bridge/Request>Resolution (Extract 2)

There are two moves in this extract. In the first move, the TL references a key phrase from the mark scheme as a form of external bridge to the concept in that shared document. The first move includes recognition from the examiner that the marking issue is resolved, as the indication of the key phrase 'reduced overproduction' is then taken as a heuristic for recognising quality in future candidate performances. In the second move, the TL expresses agreement with the examiner, therefore legitimising and validating the examiner's stated position.

In this section I have outlined how participants attain resolution; suggesting that this involved the interaction between rationale sharing, and requesting, locating and bridging information. In the next section I explore how the nature of information changed over the course of feedback interactions.

6.2.2 Case Histories of Problematic Items for Examiners

I have presented evidence of how feedback features interact to contribute to the attainment of resolution. Resolution is a signifier of effectiveness, with discourse being used to bring together alternative perspectives. Time is also a key dimension of my theoretical framework. By considering the development of discourse over a series of interactions it was possible to consider how the examiners worked to co-construct resolution around the trouble source of continued misalignment. In some instances it was evident that resolution involved a protracted interaction around some extended questions that I termed 'Problematic Items'. In these cases effectiveness was signified by the attainment of resolution over a period of time.

Analysis in the previous section (6.2.1) identified a sample of 20 items where TLs and examiners engaged in two-way discussion as part of the feedback exchange. These items were of interest because the existence of continued interaction surrounding these items indicates that the participants were engaged in an on-going process to refine their understanding of each other and the mark scheme. Of these items, four extended response items were identified where TLs and six examiners interacted beyond an initiation-response episode (Table 6.2).

Table 6.2: Extended Interaction Items and Interaction Word Count

Subject	Examiner	Year	Item	Item type
Chemistry	6	2014	7(d)	Extended
Economics	8	2014	6(b)	Extended
Economics	11	2014	6(b)	Extended
Economics	12	2014	6(b)	Extended
Geography	28	2015	8	Extended
Geography	30	2015	8	Extended

These items formed the basis of a case study analysis to explore how the participants pursued resolution over a period of time. The corpus of data for these episodes of interaction was 4225 words (a 2% sample of all of the feedback data).

Some words were used significantly more in some quartiles than in others, suggesting that they were involved in purpose driven action. For this analysis, I identified keywords that were found in a significantly greater number (i.e. signified at the 5% confidence level) in one quartile compared with the other quartiles (for an explanation of this method see Chapter 4.2.4.1). These words then formed the basis of further exploration to discern the context in which they were used. Table 6.3 shows the keywords that were found for each quartile.

Table 6.3: Keywords and Quartiles

Quartile										
1			2			3			4		
Word	(n)	%†	Word	(n)	%	Word	(n)	%	Word	(n)	%
we**	21	1.78	was*	17	1.28	think**	7	0.67	they*	9	1.59
only*	8	0.68	marginal*	6	0.45	were*	4	0.38	at*	9	1.59
						why*	4	0.38	page*	7	1.24
									gets*	6	1.06

† % of quartile corpus

* p value <0.05; ** p value <0.01

Each keyword was analysed to identify the main collocating words that related to it. From the colocation tables that were constructed in this analysis (shown in Appendix O) a series of graphical representations were constructed. These representations relate the keyword and collocating words to examples of feedback. They also include space for themes that appeared to link the word patterns to the feedback purpose.

6.2.2.1 Collocation and Quartile 1

The two keywords that were found in the first sample quartile appeared to map onto phrases that were used to enforce collective authority ([we]) and to locate credit in performances ([only]) (Figure 6.17).

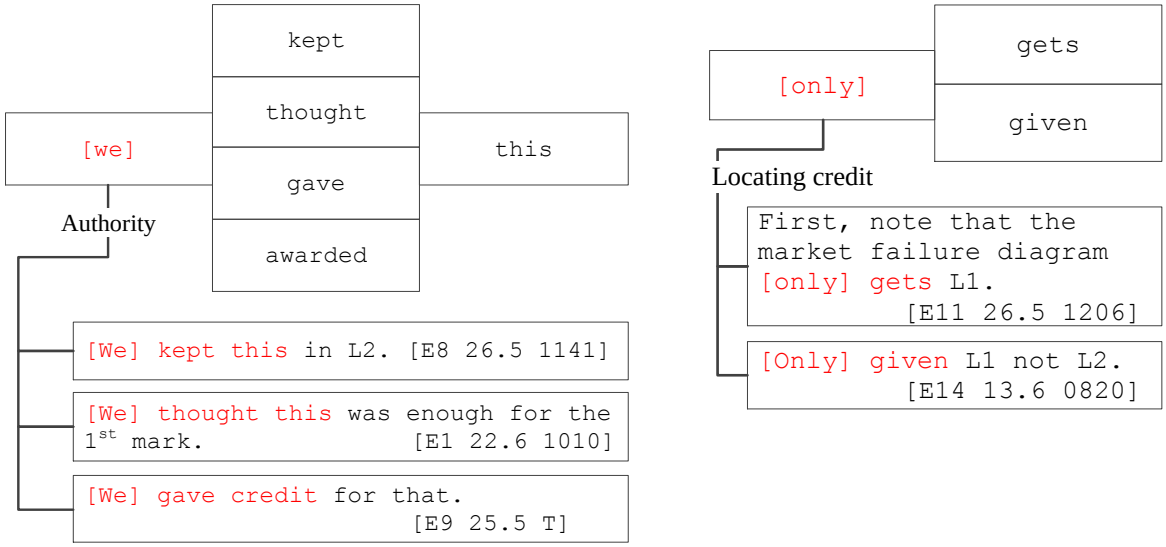

Figure 6.17: Collocation Patterns/1st Quartile

These collocation patterns reinforce the idea that the earliest stages of feedback may involve TLs using limited rationalisation. This coheres with research findings in some professional contexts that show a tendency for experts to adopt an egocentric position when giving feedback, and that this can both be a poor basis on which to base a perception of others (Derks & Bakker, 2010) and interfere with their ability to predict others' message reception (Kruger et al., 2005).

6.2.2.2 Collocation and Quartile 2

The two keywords that were found in the second sample quartile ([was] [marginal]) appear to map onto phrases used by the TL to rationalise and justify their perspective on the shared candidate performance (Figure 6.18/Figure 6.19).

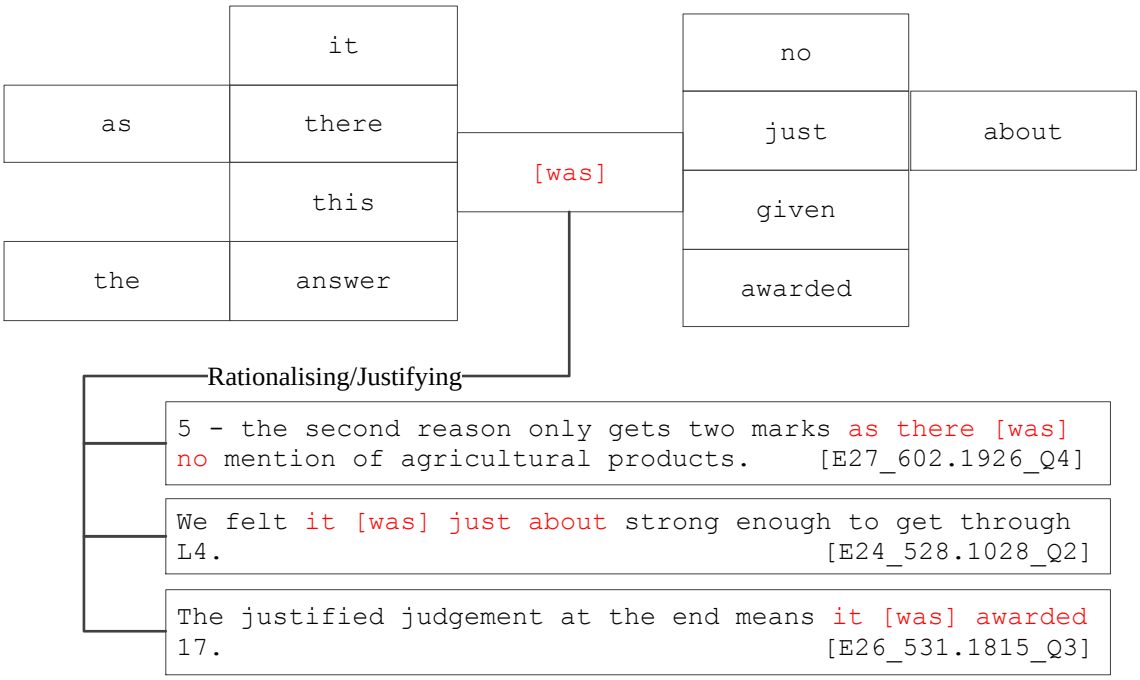

Figure 6.18: Collocation Pattern 1/Quartile 2

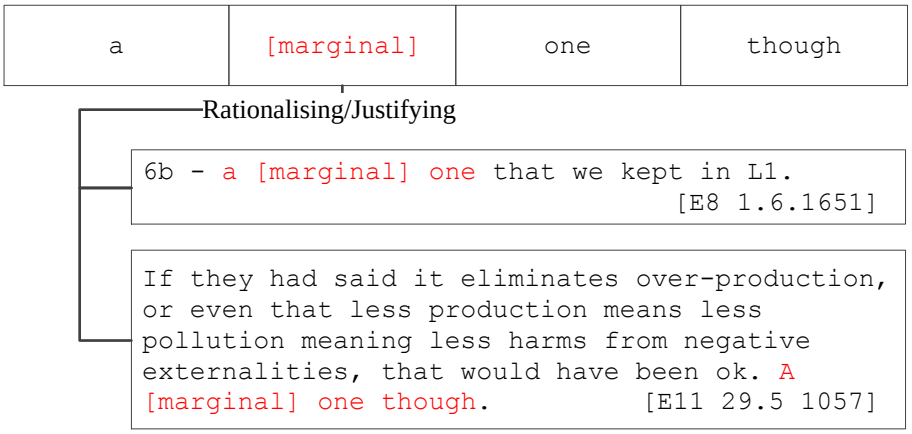

Figure 6.19: Collocation Pattern 2/Quartile 2

These patterns appeared to move the feedback discourse beyond the original interaction, which by definition failed to satisfy the needs of the examiner. In this quartile, the participants started to use language to elaborate reasoning.

6.2.2.3 Collocation and Quartile 3

The three keywords that were found in the third sample quartile ([think] [were] [why]) appear to map onto phrases that were used by the participants to continue to further

rationalise and justify their perspective, as well as to take this further and to relate their perspective to each other (Figure 6.20/Figure 6.21/Figure 6.22).

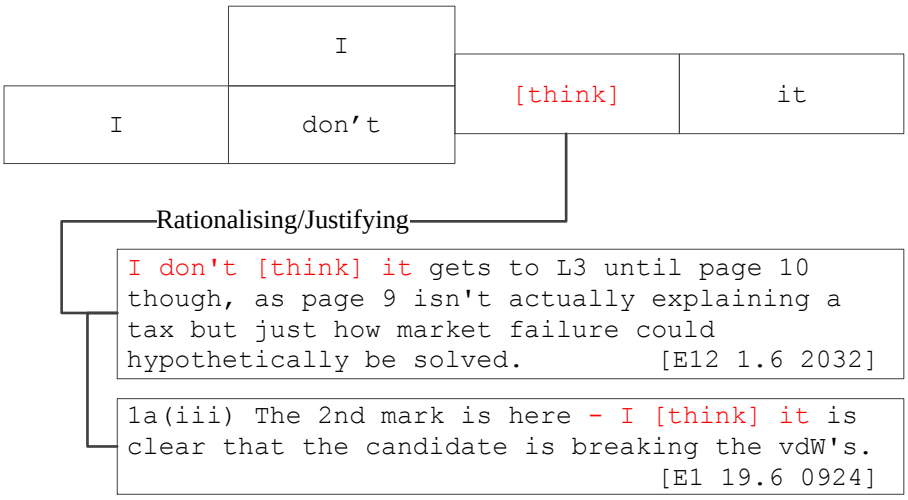

Figure 6.20: Collocation Pattern 1/Quartile 3

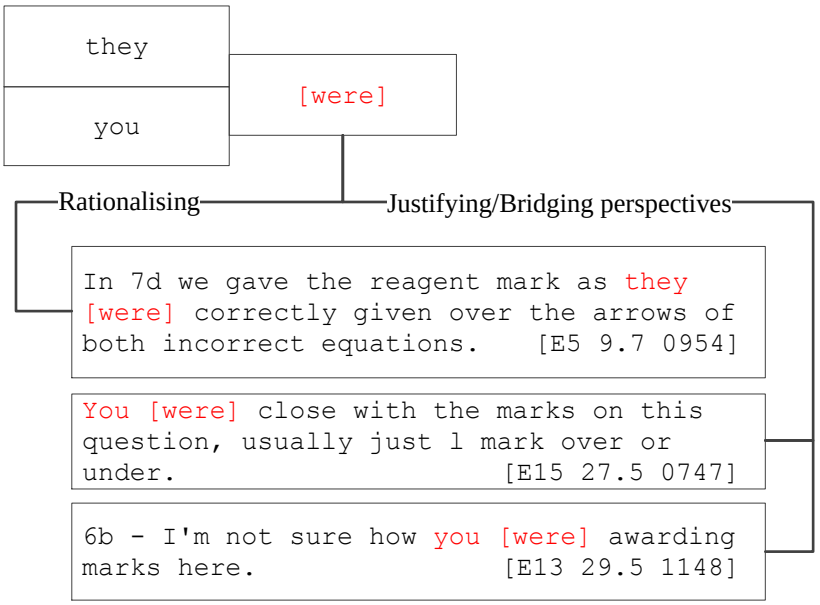

Figure 6.21: Collocation Pattern 2/Quartile 3

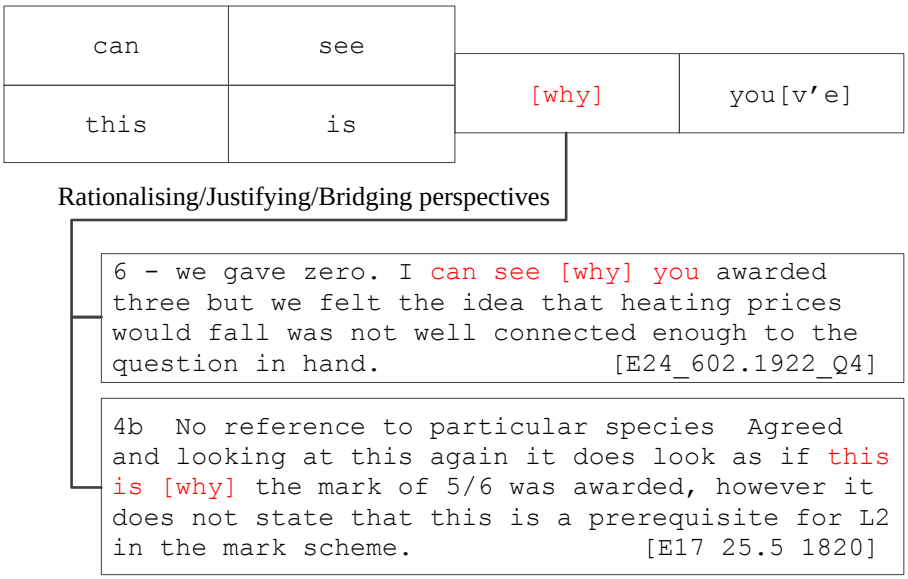

Figure 6.22: Collocation Pattern 3/Quartile 3

These patterns appear to move the feedback discourse beyond the second phase (elaboration of an individual's perspective), as this still required further interaction. During this quartile, there was evidence that the participants looked to bring together their perspective with each other in order to influence thinking.

The analyses of Quartiles 2 and 3 suggest that, taken together, these areas of discourse may share a degree of overlap. They are both characterised by the involvement of rationalising and justifying, which may be evidence of engagement in *Joint Intellectual Action* as participants decide the appropriate content to include in view of their assessment of the other's level of understanding.

Closer analysis of these feedback phases suggested that examiner rationalising, in particular, had a potentially strong influence on the attainment of resolution. Analyses suggest that when they share their own perspective examiners can (1) reflect on how their own understanding relates to an approved hierarchic position, (2) probe as yet unclear conceptual meaning and refine understandings, and (3) verify a TL's perspective as a form of reassurance. In the following extracts, I exemplify these three purposes.

1: Examiner reflects on how their understanding relates to an approved hierarchic position

Ex: I first gave this 10, then 18, then 10 I don't think tthe explanation is strong enough for band 2, bootom of page 8, top of page 9 Do you agree?

TL: This is marginal one I think it should go through Band 2 because when it says production falls, meaning less negative externalities such as pollution, meaning market failure resolved it is doing enough of the right developed analysis.
[E12 31 5 1916]

Ex: Have kept 6b at 10 as diagram is fine on page 9 - but analysis does not secure band 2 However, does the last sentence of script on page 14 secure it? Thanks

TL: A tricky one - I'd just about say no, as at the end they are saying tradable permits and tax leads to overproduction being eliminated, so it isn't saying a tax will do that alone.
[E12 9 6 1805]

2: Examiner probes meaning and refines understandings

Ex: I am still not satisfied with this [feedback]. I have looked at it again and can only assume that you wish me to credit the explanations given in lines 2-6. In these cases it is unclear as to which characteristic is being explained because the candidate makes no reference to them in these sentences.

TL: You are absolutely right that the explanations in lines 2-6 cannot be credited because it isn't clear which characteristics are being talked about. This means what you refer to as 'opposing arguments' are not really opposing arguments because no explanation marks have yet been awarded. This means it gets the two explanation marks for rival and excludable, bringing the answer up to 4/6. It then gets one comment mark at the end for reaching an explained position as to why open spaces can't be public goods (because they have elements of rivalry and exludability)
This is how the answer ends up at 5. Can you confirm you are happy with this?
[E24_524 1827_Q2]

3: Examiner reassures and/or verifies the TL

> As far as 7d goes I htought form a previous paper that the 'distil' and 'reflux' had to be wrtten as a sentence rather than just underneath the equations so I will know that for the future.
> [E1 25.6 1648]
>
> Q3. I assume it being repetitive and lacks variety means boredom so will use this in the future.
> [E8 4.6 1718]
>
> OK. I didn't think it was clear enough for Band 2, but I'll ensure that I reward it if I see 'reduced overproduction' again.
> [E11 8.6 1603]
>
> 4c Agreed. My error - focused on incorrect point about quantity demanded and ignored the most obvious correct point preceding it. Corrected to 4/6 for three extra comment marks and the analysis mark.
> [E22_527.1728_Q2]
>
> Once again this is really helpful. In particular, I think sometimes with the essays you question yourself when it's in the very low or high category, and on reflection this was a little generous, not least because of the focus on depositional landforms without enough links to erosion.
> [E30_709.0900_Q3]

It was noted earlier that sharing a perspective can overlap with requests for information. This is because a question often contains the implicit articulation of the perspective of the questioner. Requests for information are powerful because they invoke a response, which can lead to the juxtaposition of perspectives as participants respond to each other. This point is demonstrated in the following examples taken from interactions around the problematic items.

Ex: Many thanks for helpful comments. I have one query – if I understand correctly, we gave the colour change in 7d for 'orange' in 789, but not for 'green' in 860. Is this not a bit of a contradiction?

TL: We did give the colour change in 789 for orange to green (altho' the answer states the carboxylic acid went from orange to green) – probably what is seen in the reaction vessel.
[E6 20 6 1658]

Ex: I have another query for you. Can the equations *on their own* (but with distil/reflux mentioned) give *both* the equation mark *and* the 'distil/reflux produces aldehyde/carboxylic acid' mark?

TL: Yes - I am marking the distil mark and reflux mark from equations, provided the equations are correct for the aldehyde and carboxylic acid.
[E6 24.6 2053]

The invocation of a response through a request for information is a form of *Joint Intellectual Action*. A close analysis of examiners' requests for information suggests that these moves helped to attain resolution through (1) encouraging the TL to elaborate nuances of their judgement making, (2) testing their understanding of the TL's understanding, and (3) verifying that they agree with the TL. Examples of these are outlined below.

1: Examiner encourages the TL to outline a nuanced judgement

Ex: Some correct points about tax however it refers to corporation tax and implies this is why the changes are made. stick at L1?

TL: I would actually let that through L3B2 - whilst they do identify the wrong tax, it is at least a tax on firms that they have identified, and if we ignore the corporation tax part then they are right in saying indirect taxes are taxes on revenues which raise costs of production etc....just enough for BOD I think.
[E8 31.5 1910]

2: Examiner tests their understanding of the TL

> **Ex:** I know it states in the mark scheme in brackts that in need specific reference to the diagram - just to check does this therefore mean just stating price and qty increases is not enough without linking to diagram?
>
> **TL:** Yes - they must link to the diagram.
>
> [E8 30.5 1949]

3: Examiner verifies TL thinking

> **Ex:** I'm assuming with the essay that if the analysis (probably using a diagram) does not focus on the correction of market failure, I do not move into L3 Band 2? This therefore limits ability to get credit for explaining adv/disadv of tax?
>
> **TL:** Yes, that's right - an essay that doesn't hit the L3B2 criteria will effectively be capped at 10. General discussion of the advantages and disadvantages of a tax will not move the answer beyond that.
> [E11 26.5 1135]

Taken together, these findings draw attention to the mediating role of examiner rationale sharing in feedback discourse, in particular during the mid-phases of the interaction.

6.2.2.4 Collocation and Quartile 4

The four keywords that were found in the fourth sample quartile ([they] [at] [get*] [page]) appear to map onto phrases that were used by the TL to continue to rationalise their thought position with additional locational and mark reinforcement work (Figure 6.23/Figure 6.24/Figure 6.25/Figure 6.26).

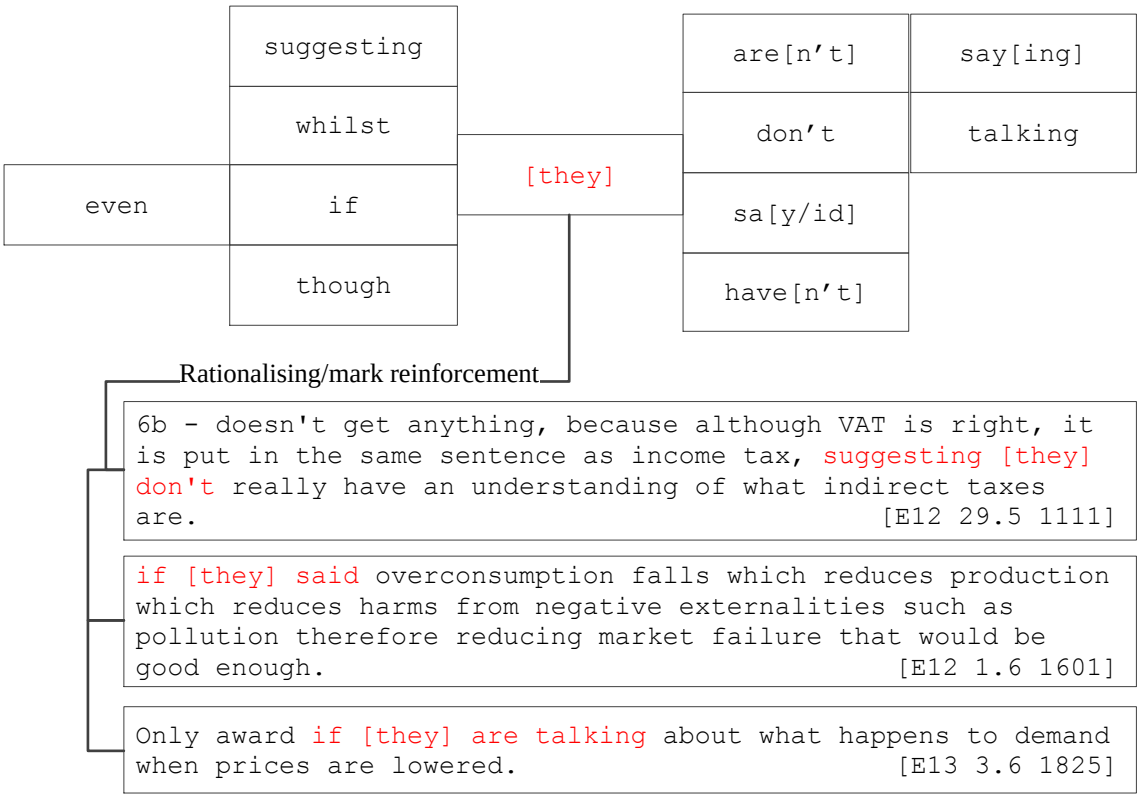

Figure 6.23: Collocation Pattern 1/Quartile 4

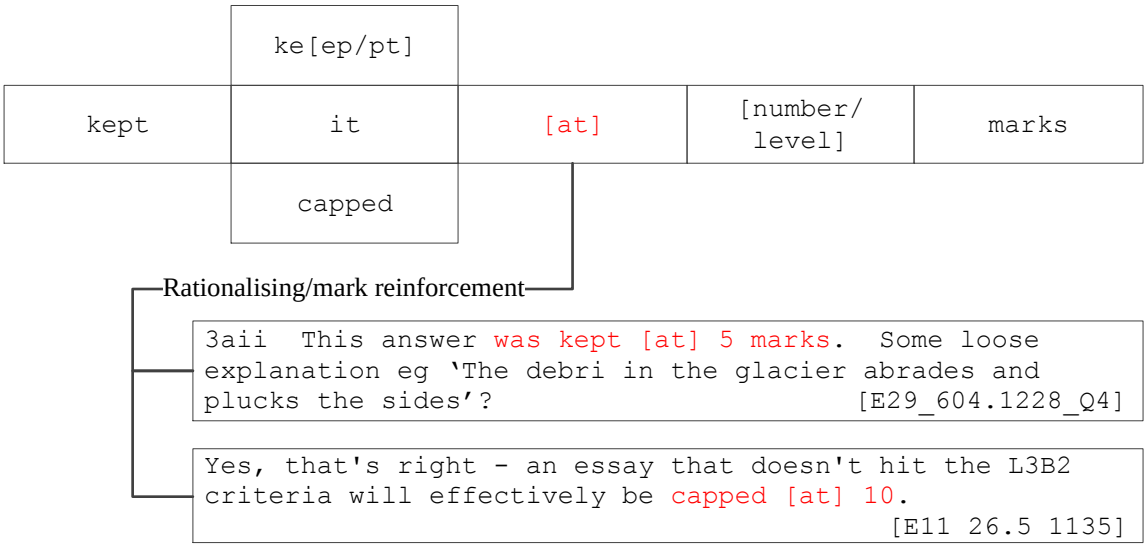

Figure 6.24: Collocation Pattern 2/Quartile 4

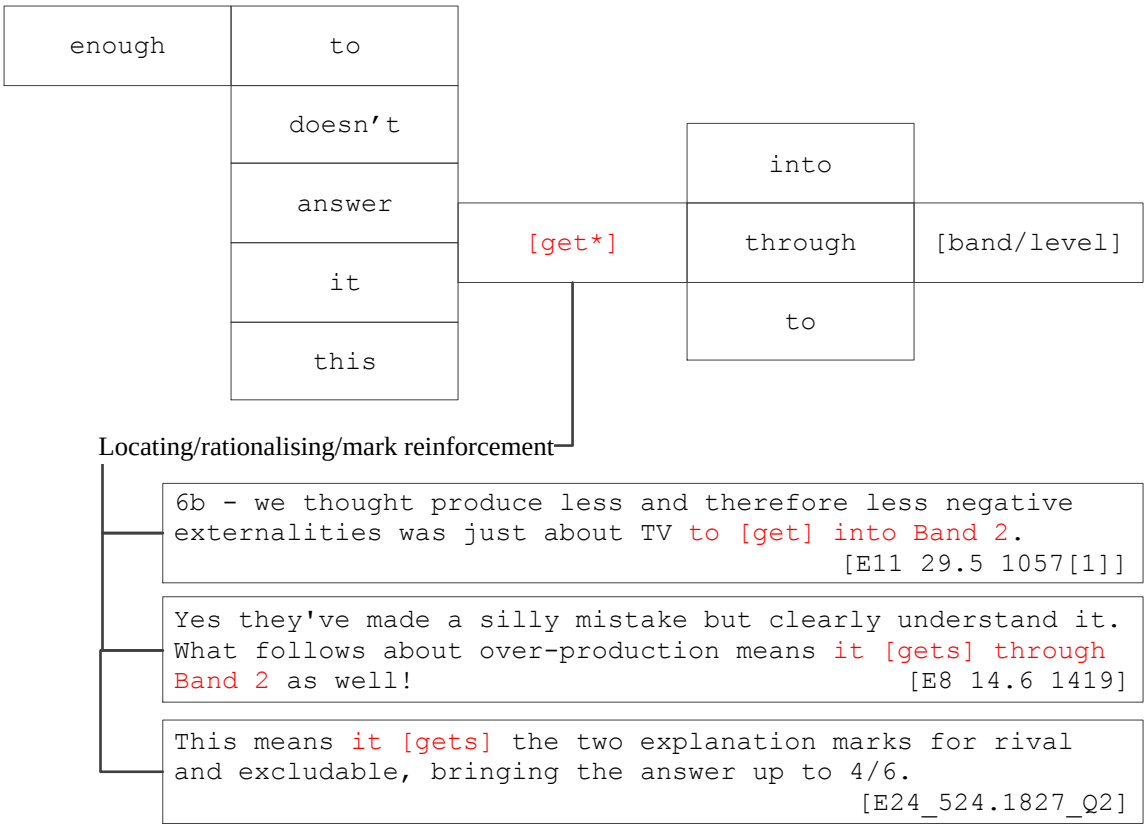

Figure 6.25: Collocation Pattern 3/Quartile 4

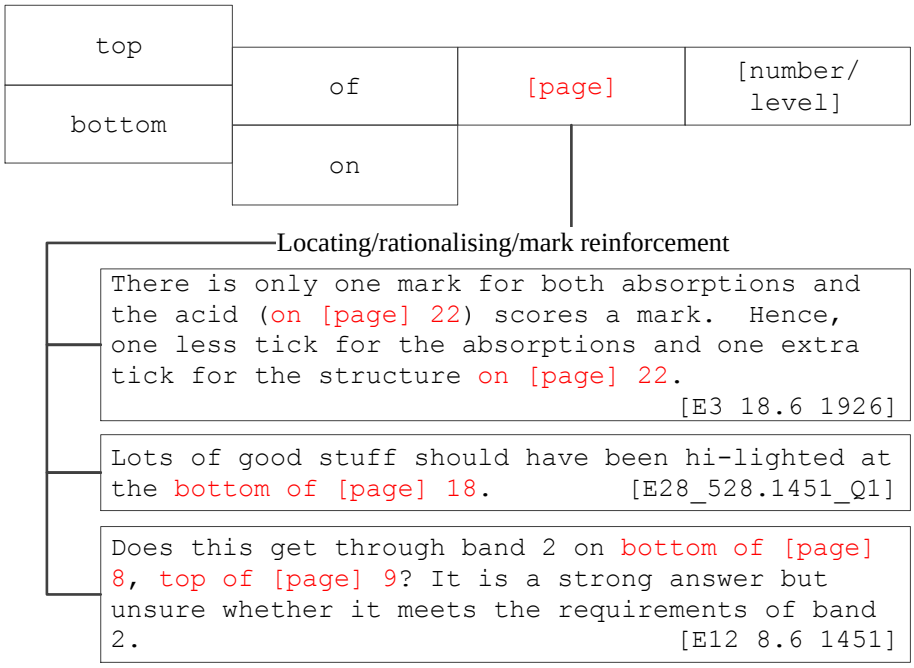

Figure 6.26: Collocation Pattern 4/Quartile 4

Although there was still space for some additional rationalisation of the TL perspective in this quartile, the analysis suggested that the feedback discourse shifted towards a conclusion by reinforcing the location of credit in the performance. This tied the discourse to the earliest quartile and acted as a cohesive device. In addition, the TL draws on the use of marks and levels as a way of further reinforcing the notion of the creditworthiness of the performance. This use of a common measurement concept further reinforced the TL's evaluation of the strength of the performance.

Below is an extract that demonstrates these elements over a protracted discussion of a problematic item over days (Figure 6.27). In the extract, a TL and an examiner discuss a 6-mark Chemistry question. In the question, the candidates need to describe the oxidation reactions of propan-1-ol when using a suitable oxidising agent, and include reagents, observations and equations in the answer. The mark scheme includes guidance that states: 'Observations: Orange to Green'. In this feedback episode, the TL refines the examiner's understanding of an unwritten nuance of the mark scheme. The analysis shows how the TL shifts from locating the creditworthy element of a performance – reinforced with implied authority [`The aldehyde equation scores a mark`], to sharing the rationale for the decision on this item [`We did give the colour change in 789 for orange to green (altho' the answer states the carboxylic acid went from orange to green) – probably what is seen in the reaction vessel. But in 860 is was just 'green' not mention of from what colour`]. The TL also relates their own perspective to that of the examiner through indicating that their marking positions are misaligned [`Your ticks and marks are way off`] before reinforcing the creditworthy elements of the candidate performance in terms of marks [`no mark around the 'gentle warming' text/it will definitely form an acid - no mark`].

From: TL: 19/06/2014 21:11
ID: 649420998:
7b NOT: Bromine water
ID: 649108490:
7d The aldehyde equation scores a mark - (an H2O below the aldehyde)
ID: 647304860:
7d The aldehyde equation isn't balanced.

>From: EXAMINER: 20/06/2014 11:08:31 AM
Hi *TL name*
Many thanks for helpful comments. I have one query – if I understand correctly, we gave the colour change in 7d for 'orange' in [script ending] 789, but not for 'green' in [script ending] 860. Is this not a bit of a contradiction?
Examiner name

>>From: TL: 20/06/2014 16:58
Hi *Examiner name*
We did give the colour change in 789 for orange to green (altho' the answer states the carboxylic acid went from orange to green) – probably what is seen in the reaction vessel.
But in 860 is was just 'green' not mention of from what colour.
TL name

>>>From: TL: 20/06/2014 17:54
ID: 64904982:
7d Your ticks and marks are way off - no tick for the last lines of text.
ID:647153093:
7d Gentle warming will definitely give the acid here - no mark around the 'gentle warming' text.
ID: 647500687:
7d As before, if you gently heat with an oxidisng agent - it will definitely form an acid - no mark.

>>>>From: TL: 22/06/2014 11:30
ID: 649193277: 7d we gave the distillation mark - the rest was perfect.
ID: 648338991: 7d the acid equation is incorrect here.

[E6 19.6 2111]

Figure 6.27: Extended Feedback Episode

6.3 Summary of My Findings for Research Question 2

In the first section of this chapter I used examples of feedback and interview data to present how effective feedback could be realised in practice. I related the concept of effectiveness to features of the IAF that I reported in Chapter 2. In particular I showed how these features contributed to the attainment of resolution through aligning alternative examiner perspectives. I suggest that this systematic alignment represents examiner learning; with examiners accruing a sense of how to interpret mark schemes in line with the expectations of more senior examiners.

It appears that resolution (which is the signifier of effective feedback) involves co-construction. Examiners learn to interpret and apply a mark scheme in line with the expectations of more senior examiners through establishing and attending to common ground in their feedback discourse.

My analyses suggest that some features contribute to this common ground building. These features include reference to shared resources, the use of clear and specific language, and the incorporation of elaborated rationales. My analyses show that mutual attention to shared references can also lead to two-way dialogue, which is perhaps the most salient indicator that resolution involves a process of co-construction where examiners become aware of each other's perspective.

According to my theoretical framework, this mutual attuning to shared references has the potential to be an Intermental Development Zone as it is a space where the TLs are aware of the needs of the examiners.

My analyses also gave insights into the trouble sources that instigate feedback. The identification of these trouble sources is important as my analysis suggests that examiners use some features of the IAF to attain resolution and mitigate the impact of some of these challenges to enable feedback to support the construction of learning. These trouble sources included:

Lack of experience and unfamiliarity

There is little evidence in the interview data of the role of experience or familiarity, which is curious given that these elements appear to be an influence on feedback (see analysis in Chapter 5). One possible explanation for this might be that these elements are phenomena

that are not salient to the examiners during their professional reflection, and are only revealed through the sort of close analysis that my methodology affords.

Assuming that experience is a proxy for aligned understanding

My analyses also give an interesting insight into how experience may represent an element of the trouble source that underpins feedback discourse. The assumption that examiners accrue experience in interpreting common terminology (through their on-going examiner work over time) was challenged by the data. Examiners explained that the meanings of common terms could change from one (annual) marking session to the next, which suggests that experience should not be taken as a proxy for expertise (in the way that some TLs appeared to do in the interview data).

Relationship management

My analyses of the feedback and interview data suggested that the construction of mark scheme understanding by an examiner is cumulative over the course of feedback discourse, as is the establishment of a relationship and trust with a TL. There was also indication in the data that face management was being attended to by the TLs in their feedback construction, and that this was a way of undermining the negative effects of disagreement. In particular, my analyses show that negative politeness is a feature of examiner discourse, with TLs using strategies to maintain a respectful distance from, and to minimise intrusion into, examiners' professional spaces. This finding runs counter to that reported in some other professional feedback situations. For example, Holmes (2001) and Holmes & Stubbe (2003) report how supervisors tend to use positive politeness strategies when aligning with subordinate colleagues.

Remote communication

My analyses also highlight some of the challenges that pertain to remote professional communication. Looking carefully at the principle findings from the literature about effective feedback (that I set out at in Chapter 2 and restated at the beginning of this chapter), my analyses support the suggestion that a lack of collocated interaction (*Mode* factor) and trust (*Source* factor) are trouble sources. Remote, text-based interaction is a challenge for effective communication because of the relatively large amount of time needed to construct information (compared with verbal communication), and because it lacks nuance and paralinguistic cues.

Mark fluidity

In addition to remote communication, the fluidity of mark decisions (due to the space for subjective judgement) makes the interpretation process challenging because examiners sometimes lack trust in the TL to give them definitive advice on whether a mark is legitimate or not. Importantly, these elements represent a challenge for the use of clear and specific language (*Language Use* and *Content* factors) in feedback, which my evidence suggests is a feature of effective communication.

To mitigate the effects of these trouble sources, my analysis suggests that examiners use a number of strategies that are evident in my IAF. The use of positive information in feedback communication (*Emotion* factor) and politeness and face management (*Language Use* factor [Distancing]) are implicit components that centre on relationship management. These elements sit alongside the explicit elements that seek to codify important information (*Knowledge* factor).

Finally, the successful mitigation of the challenges allows feedback to support the construction of learning (*Pedagogy* factor) through (1) consciously attempting to address the recipient's learning needs with the feedback messages, and (2) allowing the recipient to have an active role in the discourse (*Recipient* factor) (e.g. through enabling two-way dialogue to check meaning etc.). I interpret this use of feedback as performing articulation work (Strauss, 1985), as the use of strategies that mitigate disputation enables on-going learning and task completion.

The acknowledgement of time as a focus for analysis was a key element of my theoretical framework. My analyses of cases where TLs and examiners worked from disagreement to resolution allowed me to consider how alignment was achieved across the course of feedback discourse. In particular, the content of what the TLs and examiners communicated gave insight into the common ground that they established across their discourse. My analyses suggested that TL feedback discourse moved across four stages.

The initial stage of discourse centred on a trouble source; an indicator of disagreement that suggested that the TL and the examiners held different perspectives on a common examination script. In the first phase language was used by TLs to locate the position of credit worthy elements of a performance. Although this type of content was not generally correlated with attaining resolution, it was perhaps a logical first step in the process of

building and testing the limits of any shared common ground. The process of outlining the location of creditworthy performance could be seen as a move where the TLs sought to use shared subject content knowledge as a basis for reaching agreement. In this case, subject content (e.g. a specific concept in Economics) is a form of background common knowledge that connects the TL and the examiner to each other through their shared membership of a subject specialist community. This community is held together by the language and concepts that set them apart from other subject communities.

Following an indication of persistent disagreement, the character of feedback shifted towards having a greater emphasis on the TL sharing their rationale and elaborating on a marking point. This move represented a growth in the base of potential common ground as the TL provided the examiner with information about the content that they felt was pertinent to the disagreement.

The next stage of discourse generally involved a degree of perspective sharing, which could be instigated by either the TL or the examiner. This was perhaps the most important stage as it implicated a very rich set of interactions. In this stage the participants continued to build their common ground through relating their thinking to that of each other, and this was where the display of *Joint Intellectual Action* was most evident. TLs and examiners rationalised their thinking to each other and TLs were prone to establishing 'bridges' between information from across sources (external) or across messages (historic). This process implicated the participants in reflecting on their own thinking, engaging in refining their thinking, or reassuring each other about the convergence of their understandings.

The fourth discourse stage focused on concluding the interaction, and was a space where coherence could be reinforced. This coherence-building included references to elements of previous content, such as the location of credit in a performance, so as to reinforce the very narrow mark-based common ground that was the instigation for the interaction. This final discourse aspect appeared to act to impose meaning on the mark system through cogent and elaborated language use, in effect expanding the base of common ground associated with a mark and helping it to have a more nuanced meaning. This action accords with the sociocultural idea of learning as a process of expansion (Engeström, 1987, 2001). In this conceptualisation, learning occurs where an individual integrates the way that another person sees an object into their own concept of that object.

7. Discussion and Reflection

My study set out to explore the features of between-examiner feedback, an under-researched area of communication, and to consider the features that might make the most effective contribution to the alignment of examiners' understandings of shared mark scheme criteria. I structured my study around two related research questions.

My first research question was 'What are the characteristics of examiner feedback?', and this then allowed me to explore my second research question, 'What are the characteristics of effective examiner feedback?' I structure this chapter into two broad sections: *Findings* and *Critical Reflection on Methods, Implications, and Next Steps*.

The introduction to my thesis established the reason why the study of feedback, in general, and examiner feedback, in particular, is important. In general, I note that the behaviourist influence on the study of assessment feedback has led to a downplaying of complexity in the conceptualisation of feedback-giving practice. In particular, I argue that examiner feedback is an important area of study because of the link between examiner communication and marking reliability. In high stakes examination situations, feedback is expected to support the alignment of between-examiner thinking and action. Therefore, understanding the nature of communication allows insights into the nature of the problem of marker reliability, and how this can be remediated through best feedback practice.

My adoption of a sociocultural analytical framework helped me to move my study beyond a focus on the content and directionality of feedback communication (a hallmark of some previous feedback studies), and to consider some of the contextual features that influence examiner feedback.

7.1 Findings

This discussion is organised into two sections. The first section discusses the characteristics of examiner feedback that I uncovered through my close analysis, whilst the second section discusses the concept of effectiveness, and in particular how resolution was attained through feedback in relation to these features.

7.1.1 Research Question 1: What are the Characteristics of Examiner Feedback?

The first objective of my study was to explore the characteristics of examiner feedback. Based on the sociocultural theoretical framework that I adopted for this study I argued that examiner feedback is a professional learning discourse, and I anticipated that feedback would

act as a repair mechanism that centred on a trouble source (which would principally be the existence of weak mutual common ground between examiners). My theoretical framework also suggested that common ground establishment would enable the construction of an Intermental Development Zone (IDZ), where participants' minds are mutually attuned.

Finally, drawing on theoretical perspectives beyond conventional sociocultural theories I argued that common ground construction would involve some form of 'articulation work' (Strauss, 1985) on the part of the participants to ensure that their communication achieved its coordinating function (and that marking tasks were completed).

To identify the characteristics of examiner feedback I developed a methodology (drawing on and extending the methods employed by other studies that have adopted a sociocultural approach to analysing communication (e.g. Mercer, 2004; Littleton & Mercer, 2013)). This augmented Sociocultural Discourse Analysis (ASCDA) methodology, allowed me to investigate the features of interaction at both a particular and a general level, and clustered my analysis into four specific feedback discourse themes: feedback *Content*, the development of discourse over *Time*, evidence of *Joint Intellectual Action* within feedback, and the *Impact* of feedback.

My analyses identified three broad characteristics of examiner feedback: it focuses on a trouble source; it is used to convey a good deal of negative information (disagreement); and, its content has both transactional and interactional elements (which are heavily interconnected).

7.1.1.1 Feedback is Discourse that is focused on a Trouble Source

My theoretical framework anticipated that feedback would act as a repair mechanism and centre on a trouble source, such as the existence of weak mutual common ground between examiners. My analyses suggest that these trouble sources could be found in three areas which related to: the subjective nature of marking; limits around examiners' shared experiences; and, the developmental nature of learning.

The Subjective Nature of Marking

My analyses offer evidence to support the idea that subjective, levels-based marked items are a trouble source, with more feedback being given for those examination papers that include the most subjective items (i.e. Geography and Economics). This observation links with ontological considerations about the nature of the mark that is expected to form the common

ground between the examiners. There is an assumption that predetermined monitoring marks are fixed in character. This assumption of stability (that a performance warrants a single agreed mark that can be recognised by professionally trained examiners) is perhaps confounded when it is set within an interactional environment. The feedback environment creates an opportunity for a TL to gain additional insight into how an examiner sees a performance, and this has the potential to shift the TL's original judgement of a performance. This element of reverse hierarchic influence has been noted elsewhere (Johnson & Black, 2012b), and suggests that feedback interaction has the potential to undermine the ontological assumptions of mark stability that underpin the marking hierarchy.

Limits around Examiners' Shared Experiences

A second trouble source was anticipated to be any lack of shared experience between the examiners, since this would indicate that they lacked resources of past experience to draw on in their practice (and would require some communication work to remediate). My analyses offer support for the notion that feedback was functioning to align perspectives and diminish common ground weaknesses with more content being communicated to new and unfamiliar examiners (than to experienced and familiar examiners).

The Developmental Nature of Learning

This links with the final, and perhaps the most important, trouble source which was that feedback content was crafted to cater for the learning needs of the recipient examiner. My theoretical perspective suggested that common ground establishment, which includes work related to the construction of an Intermental Development Zone (IDZ), would be most pronounced at the earliest stages of feedback interaction. My analyses offer support for this, with most feedback being found at the initial stages of communication (where it would be anticipated that learning needs would be most pronounced, and shared common ground most weak). One manifestation of the developmental nature of feedback was in how its content changed over the course of the marking period. In the earliest stages there was a suggestion that there was more directive activity, and that the examiners engaged with the reasoning for any differences of perspective later in the marking discourse (as trouble sources around specific marking issues appeared).

7.1.1.2 Disagreement is a Significant Feature of Feedback Content

A consequence of trouble sources being a focus for feedback discourse is that feedback content tends to be heavily infused with disagreement. This point (and its consequences for articulation work) is discussed in more detail below.

7.1.1.3 Feedback Content has both Transactional and Interactional Elements (which are Heavily Interconnected)

Transactional Feedback Qualities

My analyses showed that feedback content had a transactional quality which communicated the content that supported the professional development of the examiner (i.e. the information that better helped them to understand how to think within the examiner community).

Identifying the location of credit and rationalising why credit was given were common types of information given via feedback. This observation suggested that helping examiners to understand where and why credit should be given was central to the development of an examiner's developing practice.

According to my theoretical framework, the judicious choice of content to include in feedback communication represented a form of Joint Intellectual Action (JIA). Language choice signified where a feedback giver was considering the perceived needs of a recipient examiner, and I argue that this represents a form of mutual attuning because it indicates that the participants have a sense of each other when they choose their form of communication. My analysis suggested that the examiners accomplished these acts in a variety of ways, which I was able to cluster into *Bridging* themes. These themes included where examiners made links between relevant information sources (e.g. mark schemes and performances); where examiners used questions to prompt reflection; and where the examiners contrasted their own practice with that of the other examiner.

Finally, examiners were also given directive information which acted as a way of connecting the examiner to the conventional practices that helped to define the examiner community (e.g. the ways of doing things through the digital marking tools).

Interactional Feedback Qualities

My analyses suggested that feedback was performing a relationship management function, with communication style potentially reflecting concerns about the potential effect of the

message on the recipient. This point links with the observation (outlined above) that feedback communication was heavily infused with disagreement. In this way, and in line with observations by Brown & Yule (1983) about language in general, feedback was performing an interactional (as well as a transactional) function.

Some of the stylistic elements of feedback construction were evident in the way that examiners dealt with the 'within message' chronology of feedback construction. The construction of message opening and closing structures were purpose driven, with message openings being the location of interpersonal distancing strategies that included indications of politeness. This was especially the case for the most subjective items, for new and unfamiliar examiners, and particularly given the prominence of negative information content in the feedback messages as a whole.

Based on my theoretical perspective I argue that the prevalence of negative discourse has implications for professional face management, and that the examiners would be expected to engage in mitigation strategies (i.e. articulation work) to maintain positive professional relationships. My theoretical framework anticipated this issue, drawing on the work of Goffman (1967), who argues that the presence of negative information in social interaction embroils issues of face management. I argue that the use of negative politeness, shifting the focus of discourse from strong (potential) disagreement towards weak disagreement, and underplaying differences, is a form of articulation work. This hidden coordination work helps to maintain positive professional relationships through refocusing the discourse from a disputational to a cumulative discourse. This cumulative discourse is important as it enables the examiners to continue to engage with differences in perspectives and to develop their common ground through bridging perspectives (e.g. a shared focus on resources).

7.1.2 Research Question 2: What are the Characteristics of Effective Examiner Feedback?

My analyses looked at how examiners moved towards alignment through their discourse, which I interpreted as being an important developmental accomplishment.

I considered effective feedback to be communication which aligns examiners' perspectives through the building or maintenance of common ground, and which therefore supports the attainment of examiner resolution. For my study I explored effectiveness through two approaches. My first approach was theoretically driven drawing on empirical evidence from observations and stimulated recall interviews with examiners, and linking the outcomes of

analysis to my Integrated Analytical Framework (IAF) to show how resolution was attained. My second approach was problem-oriented and empirically driven, where I analysed a sample of case studies where either alignment was signalled or where the pursuit of agreement was being evidenced over a series of interactions. Both of these approaches relied on the analyses that I generated through my ACSDA approach (that focused on elements of *Content*, *Time*, *Joint Intellectual Action*, and *Impact*).

In each of the following analyses I demonstrate how examiners attained resolution through their reference to shared resources (e.g. common texts and conceptual content) that were introduced into the discourse at pertinent occasions. I took decisions to introduce such references as being *Joint Intellectual Action*, which is the notion that a participant chooses communication content and structure based on a construction of the perceived needs of the recipient. These needs also included elements of relationship management. I took resolution as my measure of effectiveness (i.e. accomplishing the goal of reducing misalignment and enabling an examiner to understand a senior examiner's interpretative position).

7.1.2.1 Theoretically Driven Approach

Considering my first analytical approach, my analysis of the examiners' perspectives on effectiveness was facilitated by linking them to theory, so as to see how their views related to findings in the broader literature. To do this I related the examiner interview data to my IAF to consider how the resolution of alternative perspectives was attained. The outcomes from this analysis showed that a number of theorised elements characterised the effectiveness of examiner feedback. These analyses focused most heavily on the *Content* and *Time* elements of my sociocultural analysis.

The analyses supported the idea that effective feedback discourse enabled the construction of learning. Feedback communication was a site where differing personal perspectives were brought together and this formed a resource for common ground building and the attainment of resolution. This construction made use of dialogue (iterative interaction) where the participants could indicate areas of insecure understanding or see a problem from another perspective (e.g. Edwards, 2012). Some findings indicated that superficial learning occurred where feedback failed to encourage the active involvement of the examiner in sense making.

Alignment was maintained through reassuring and frequent communication, which contributed to an on-going and cumulative interaction. This supported the idea that

immediate feedback was a feature of effectiveness. Clear and specific language use, including specific referencing and the use of shared or recycled terms, also aided meaning alignment.

The IAF also anticipated a number of issues that would be problematic for on-going examiner communication, and which feedback would need to alleviate through the inclusion of features that could facilitate resolution. Alongside the principal trouble source that spurs examiner feedback (misaligned marks), elements of communication mode and trust would also represent elements of 'communication trouble'.

Remote, text-based interaction (i.e. a lack of collocated interaction) was a challenge for effective communication because of the relatively large amount of time needed to construct information (compared with verbal communication), and because it lacked nuance and paralinguistic cues. Evidence suggested that information that was conveyed through the richer communication channels (e.g. telephone) positively supported common ground building.

The fluidity of mark decisions (due to the space for subjective judgement) also made the feedback interpretation process challenging for examiners. Whilst TLs used their feedback messages to construct a relationship of trust with examiners (through the conveyance of useful information), where feedback lacked rationalisation it was possible for examiners to perceive a basis of distrust. A lack of clear and specific language use in feedback communication therefore represented a challenge for trust building.

Finally, feedback dealt with multiple functions, including elements of relationship management and considered the situated needs of the recipient examiner. These dual purposes represented a key component of feedback complexity and implicated a central tension of feedback giving practice, which was the management of social relations whilst also delivering honest performance feedback.

Given these tensions, it appeared that effective feedback involved a number of mitigation strategies that allowed on-going communication to flourish (and therefore support the completion of the marking tasks through enabling alignment). I conceptualise these strategies as being elements of articulation work (e.g. Strauss, 1985), as they were not primary features of the feedback concept at the outset of this study and could therefore be considered as 'hidden' practices.

Politeness and face management appeared to factor into decisions about how to craft feedback communication, and these could be considered to be elements of effectiveness. Examiners' ambitions to use clear, unambiguous language sometimes created a tension with their desire to create greater social cohesion through language (i.e. to maintain on-going engagement and to overcome the consequences of accrued negative interaction).

Drawing on Strauss' (1985) concept, I observed that TLs were very conscious of the need to ensure that competent examiners (or those who the TL believes will become so) are motivated to complete their marking tasks. The consequence of losing competent examiners from the workforce before marking is completed represents additional workload for the TL, and a concomitant strain on their relations with other examiners who then share the additional workload. The incorporation of politeness into communication is a strategy that can be used to build social cohesion between participants (Brown & Levinson, 1987).

Morand (2000) notes that criticising, disagreeing and interrupting represent potential face threat, and that the effects of this can be minimised by the use of negative politeness strategies, which are a form of 'redressive facework' that reduce friction (Culpeper & Haugh, 2014; Lakoff, 1979). It has been observed elsewhere that when superiors use politeness they tend to employ positive politeness tactics (e.g. cueing perceptions of 'nearness' and friendship), particularly with new and unfamiliar subordinates (Holmes, 2001; Holmes & Stubbe, 2003). My observations suggest that examiner feedback runs counter to this observation, with TLs commonly displaying negative politeness in their interactions with examiners (e.g. reducing perceptions of intrusion and maintaining a respectful professional distance).

I also observed that examiners incorporated positive information in feedback messages. This was perceived as being motivational and was more prone to being well-received, compared with negative content that had an adverse emotional impact on examiners and left them feeling anxious.

7.1.2.2 Problem-Oriented Approach

For the final section of my study I carried out two forms of analysis, one that was problem-oriented, and another that was empirically driven. In the first analysis I used qualitative methods to explore a sample of case studies where alignment was signalled. For the second analysis I used qualitative methods to explore the pursuit of agreement over a series of interactions. These analyses drew on all of the elements of my sociocultural analysis, but

referenced in particular, the elements of *Joint Intellectual Action* and *Impact* (i.e. evidencing how the examiners were working together to attain resolution and agreement).

Resolution Cases Studies

My analyses suggested that there were a small number of patterns that could describe how Giving Information (*Content*) interconnected with Bridging (*Joint Intellectual Action*) to attain Resolution (*Impact*). Three forms of information giving were present in the episodes where resolution was attained: *Rationalising*; *Requesting*; and, *Locating*. Furthermore, the power of this information was increased by the way that examiners bridged the information to concepts located in external sources, such as shared mark schemes or examination scripts. The bridging process was an indication that the participants were drawing on resources to reinforce the key embedded ideas and concepts pertinent to the marking task. This analysis coheres with the theory that informs my study, which is that the process of productive shared development is underpinned by a mechanism of establishing and maintaining common ground through discourse (Edwards & Mercer, 1987; Horton & Keysar, 1996; Littleton & Mercer, 2013).

Problematic Item Case Studies

The development of discourse over time was a key feature of my theoretical framework. Looking at the feedback cases where TLs and examiners worked towards alignment with each other there is evidence that the TLs mediated how discourse was structured so that examiners were engaged in the construction of meaning over time. This gives insight into how the examiners used different feedback features that were responsive to each other's needs in order to overcome disagreement (and to attain resolution) at different stages of their interaction. My analyses indicated that feedback discourse moved across four stages, with most feedback being given in the initial stage.

As has been stated earlier, analyses showed that indicators of disagreement were highly prevalent in feedback communication, suggesting that the TLs and examiners held different perspectives on a common examination script. This was unsurprising since the instigation for the discourse was likely to have been an indication of a marking discrepancy between the examiner mark on a script and the pre-agreed mark for the same script. A mark indication is also a very limited basis upon which a TL could decide how best to approach feedback-giving in order to develop the examiner's mark scheme understanding.

Given the time-pressured nature of the TL monitoring and feedback activity it was not surprising that TLs would employ the strategy of identifying the location of credit without additional rationale sharing when they first noticed misalignment with an examiner. Reference to shared background common knowledge through, for example, the use of words as shorthand objects that encapsulated elaborate conceptual subject knowledge, is a quick way of drawing an examiner's attention to where the concept is applied (in the TL's view) and reinforces their shared background common knowledge with the TL. For this strategy to be successful the examiner needed to infer from the content and location information about why a performance element was credit worthy. Continued misalignment signified that the TL's and the examiner's background common knowledge was not yet strong enough to support shared coordinated understanding.

Some of the problematic issues relating to language interpretation have been outlined elsewhere (e.g. Wiliam, 1993), with this implicating the semiotic qualities of words (e.g. Lotman, 1988). The problem for assessment is that differences in language interpretation contribute to inconsistent mark scheme use from one examination session to another, and this reinforces the need for intensive standardisation activities prior to each examination session. It is not enough to employ examiners from one session to the next without standardisation because marking is a complex activity that is prone to shifts in the conceptual base of linguistic interpretation over time. This point also acts as a reminder that professional activities (such as examining) are prone to the inclusion of elements of tacit knowledge. As well as defying transparent articulation (Eraut, 2000), the codification and communication of important knowledge across the examiner community is challenged by remote communication through media-weak channels (which I have conceptualised earlier as an additional trouble source). The limitations of language to convey unambiguous meaning (expressed above) are compounded by a QA system that divides examiner labour (so that all of the examiners cannot interact with each other at the mark scheme definition stage), and stipulates a communication system that encourages mainly text-based interaction.

When misalignment persisted, despite appeals to the application of information from subject specific background common knowledge, the TLs had evidence that there were areas of examiner content and understanding that needed to be developed. It was from this point that feedback linked to the idea of examiner development and the conscious construction of an Intermental Development Zone (IDZ). Where continued disagreement persisted a TL needed to use feedback to create dynamic common knowledge. They did this through situating the

specific features of the examination script performance alongside an elaboration of their particular interpretation of the mark scheme, and this was a characteristic of the second stage of feedback discourse.

This information situating process represented a growth in the base of potential common ground as the TL provided the examiner with information about the content that they felt was pertinent to the disagreement. This element also represented a point in the discourse where the TL controlled the focus of the communication so as to enable the examiner to better understand the recognised, hierarchic way of understanding and applying the mark scheme. This information could take the form of restating and reinforcing a point by using the wording of the mark scheme so as to draw on the participants' common subject language (and the implied concepts).

The third discourse stage generally involved a degree of perspective sharing, which could be instigated by either the TL or the examiner. This was perhaps the most important stage as it implicated a very rich set of interactions. In this stage the participants continued to build their common ground through relating their thinking to that of each other, and this was where the display of *Joint Intellectual Action* was most evident. TLs and examiners rationalised their thinking to each other and TLs were prone to establishing 'bridges' between information from across sources (external) or across messages (historic). This process implicated the participants in reflecting on their own thinking, engaging in refining their thinking, or reassuring each other about the convergence of their understandings.

This third discourse stage also highlighted the two-way nature of the feedback that contributed to the learning and development process. Despite being a potential threat to the implicit top-down assumption of the hierarchic marking model, two-way interaction appeared to provide an important function in the process of forging common ground. TLs and examiners engaged in requests for information from each other that appeared to satisfy both interactional and developmental purposes. At times, TLs and examiners requested further information from each other, and this information could go on to become a resource for further interaction. On the other hand, some requests were rhetorical, and were primarily made by TLs to spur an examiner to self-reflect on their thinking. These requests carried an implicit steer as they expected the examiner to interrogate the basis underlying their difference in understanding from that of the TL.

The fourth discourse stage focused on concluding the interaction, and was a space where coherence could be reinforced. This coherence-building included references to elements of previous content, such as the location of credit in a performance, so as to reinforce the very narrow mark-based common ground that was the instigation for the interaction. This final discourse aspect appeared to act to impose meaning on the mark system through cogent and elaborated language use, in effect expanding the base of common ground associated with a mark and helping it to have a more nuanced meaning. This action accords with the sociocultural idea of learning as a process of expansion (Engeström, 1987, 2001). In this conceptualisation, learning occurs where an individual integrates the way that another person sees an object into their own concept of that object.

In summary, my analyses outline some of the components that characterise examiner feedback (my first Research Question), which was a previously hidden practice. Drawing on my theoretical framework, the characteristic features of feedback included transactional content (such as information giving) which often coalesced around disagreement. My analyses also showed that the examiners drew on references to things beyond their immediate context (Bridging). My analyses also demonstrated some differences in the volume of feedback, which were indicative of the trouble sources that inspired feedback discourse. Although the effect sizes were limited, there were nevertheless some significant differences in the quantity of feedback given to examiners based on the degree of marking complexity or on the degree of shared common ground that already existed between the examiners. Finally, my analyses offered insights into how examiners employed politeness in their discourse, suggesting relationship management to be an interactional achievement of feedback. I claim that this practice forms an element of articulation work.

My analyses outline the characteristics of effective examiner feedback (my second Research Question). To answer this question I drew links between my gathered observations and established theory (by relating the perspectives of examiners to my IAF), and through considering the direct empirical links between feedback data and examiner alignment (i.e. resolution). Taken together, my analyses of how examiners move towards alignment through their discourse evidenced how the examiners achieved an important developmental accomplishment. This process, (which is by definition an effective process, given that my concept of effective feedback is one where alignment is achieved), was characterised by the use of discourse to enable an examiner to construct an understanding of a TL's marking perspective. This movement involved reference to shared resources (e.g. common texts and

conceptual content) that are introduced into the discourse at pertinent occasions. Decisions to introduce such references were based on *Joint Intellectual Action*, which I took to be the notion that a participant chooses their communication content and structure based on a construction of the perceived needs of the recipient. My analyses suggested that these occasions were triggered by a trouble source (such as mark disagreement), and that the potential disruption to developmental discourse was mitigated by articulation work which employed politeness strategies. Viewed from a broader context, this movement also represented a process where an examiner comes to be recognised as a competent professional whose judgement-making coheres with that of a recognised authority (senior examiners). This accomplishment is characteristic of a professional induction process, with the move from disagreement to alignment mirroring the description by Lave & Wenger (1991) of how a professional moves from the periphery to the centre of a community of practice.

7.2 Critical Reflection on Methods, Implications and Next Steps

7.2.1 Contribution to Knowledge

My study makes a number of contributions to the field of knowledge. One important contribution is the establishment of conceptual relevance. In this study I successfully applied the concept of SCDA to a new context; the field of remote, asynchronous, professional adult learning. This contrasts with the face-to-face, synchronous classroom-based environments in which the concept has been previously applied (e.g. Mercer, 2000; 2002; 2004; 2008a; 2008b; Mercer & Littleton, 2007; Mercer, Littleton & Wegerif, 2004). The study findings give insight into how remote, asynchronous, professional communication

In addition, my study also makes a methodological contribution through the way that it builds on the themes that are found at the core of Mercer's (2004) Sociocultural Discourse Analysis (SCDA). My approach was linked to Mercer's work through the common themes that I located within the original SCDA approach (*Content, Time, Joint Intellectual Action,* and, *Impact*), and these were placed at the heart of my analysis.

The development of an augmented Sociocultural Discourse Analysis (ASCDA) involved my consideration of the specific contextual features of TL feedback and to consider carefully the relationship between methods and my research aims (i.e. to simultaneously identify the characteristic features of examiner feedback as well as to explore the features that contributed to its effectiveness in aligning the perspectives of the participants).

In the context of a series of interactions that spread over a period of time, I needed to develop a methodology that could evidence changes in content over a large body of interaction data. The thematic focus of my ASCDA approach led me to identify a number of methodological tools that I could use to explore the nature of examiner feedback. These tools drew on methods that were located in the traditions of Thematic Content Analysis (TCA), Conversation Analysis (CA), and Corpus Linguistics (CL). For an in depth discussion of this integration process, see Johnson (2016d).

It soon became evident in my approach to coding analysis that Content involved both transactional (i.e. information giving) and interactional (i.e. interpersonal focus) elements. These elements mirrored the observation of Brown & Yule (1983). In conjunction with the Bridging and Action codes my coding schedule for analysing the data, which I validated through double coding, constructed a general picture of examiner feedback that I could use as a basis for identifying and exploring any patterns in the data (this is discussed further below) and to relate these observation to established sociocultural theory.

My theoretical framework also required me to include direct evidence of message intention and reception in the analytical process, so as to consider how anticipated indicators of feedback effectiveness mapped onto empirical interview data. This reflected the situated nature of intention and reception on how meaning is constructed through discourse (i.e. that there can be contextual features that are pertinent to a participants' experience that influences their understanding of discourse).

To support this process I used observation and stimulated recall interviews, but my concerns about how dependable the examiners' reported perspectives (e.g. Calderhead, 1981; Lee, Landkin, & Carter, 1992) led me to include a number of measures.

I carried out all of the interviews as close as possible to the time of feedback receipt. This helped to ensure that the memory trace of their initial perspective on the feedback was still accessible. To facilitate the interview process, I provided all examiners with an advanced copy of the feedback that was to be the focus of the interview.

Interviewer variance is a concern for critics of interview methodology, and this is an issue that relates to the social dynamics inherent to the interview setting. I sought to reduce this risk through imposing a standardised schedule on my data gathering interactions with examiners, and using a semi-structured interview format.

One concern relating to interview structuring is that there is a trade-off between systematic researcher control via interview structuring, and eliciting unanticipated, but valid, recipient responses (Frankfort-Nachmias & Nachmias, 1996, p.237). This potential methodological weakness can lead to unexplored features and gaps in evidence that influence the inferences made during the data analysis stage (for more on this see the limitations section below).

The insights gained from the integration of these approaches enabled me to focus on the interaction of content (transactional) and the interpersonal (interactional) dimensions of discourse, and this second element is outlined in more detail below. Drawing on a variety of methodological tools I was able to analyse content across a large body of data, to consider how the shape of communication changed over time (a proxy of development), and to explore statistical interaction between aspects of the feedback communication. In conjunction, these methods allowed me to develop the concept of articulation work to describe the taken-for-granted interaction work that feedback givers engage in to facilitate the development of the feedback recipients (for more on this see below).

I feel that this approach was successful as a method for making sense of a large data set whilst also retaining a sense of the context in which discourse was being formed and used. The trade-off between scale and detail is often cited as a weakness of specific approaches such as CL or CA (e.g. Billig, 1999). Through integrating these methods (around their common focus on how they give insights into the process of social action through discourse), I was able to retain the sense of how discourse was being used in particular situations, whilst also constructing a broader view of the discourse that was generalizable beyond a single instance.

This integration of methods also allowed me to stretch the SCDA approach beyond its traditional basis in classroom talk, and to demonstrate how development can be evidenced in text-based interaction. My ASCDA approach demonstrated how the analysis of change in discourse could be carried out across episodes of asynchronous communication. This was, again, a departure from the synchronous classroom interaction that was the original focus of SCDA. These developments are important because technological advances in communications technology mean that the development of professional expertise in large organisations is increasingly remote in nature. This makes my methodology relevant to other hierarchically and remotely organised professional environments, such as healthcare, police, military and aviation environments to name a few (Johnson, 2016c).

A significant insight that was generated by my study was how relationships were managed through discourse. The role of politeness (distancing) in the development of aligned thinking was not anticipated by traditional SCDA, but my analyses were supported by ethnomethodological theory that considers the role of face threat and its mitigation through discourse (e.g. Goffman, 1967). By involving the interactional implications of ethnomethodology in my approach I was able to construct a model of discourse development that described how examiners avoided the consequences of disputational discourse (Mercer, 2000) and maximised the potential for the construction of an Intermental Development Zone (IDZ). In the IDZ the examiners were able attune their thinking through bridging and to share perspectives. I argue that the use of politeness (distancing) performed articulation work through enabling coordinated work across individuals.

Finally, my theoretical framework foregrounded the concept of *Time* in interaction, and this contributed to my construction of effective examiner feedback in a significant way. Changes in the features of interaction over time (in my case, the Quartiles of feedback) allowed me to infer that the participants' understandings of common concepts and of each other were developing (as would be anticipated during the process of building common ground).

As well as considering 'across message' time I was also able to consider time in a more limited sense, namely, 'within message' time. My coding process allowed me to sample specific instances of feedback that demonstrated resolution, and to explore the process of attaining alignment in some depth. One benefit of this analysis was that it allowed me to identify patterns in message construction, and, for example, to consider how elements of opening and closing were purpose driven and implicated recipients' needs.

In addition, this approach allowed me insight into the role of politeness (distancing), and in particular the use of negative politeness and articulation work. I was able to note that the use of negative politeness by superior professional colleagues distinguished the examiner feedback context from other apparently similar environments (c.f. Holmes, 2001; Holmes & Stubbe, 2003).

7.2.2 Limitations

There are a number of limitations that relate to the methodology that I have constructed. The most significant limitation is that my ASCDA approach is mainly suited to the study of dynamic common knowledge. The existence of this form of knowledge becomes evident to the researcher as it emerges in the course of, and is demonstrated within, interaction. It is

apparent that this is only one element of the common ground that participants in interaction draw upon when engaging in interaction.

Examiners share a base of common knowledge that pertains to their membership of a subject specialist community. As I outlined earlier (Chapter 6), this presents a problem for analysis as such background knowledge exists as 'latent content' in the communication that takes place between examiners. This latent content is a challenge for analysis as it is only through the participants' implied allusion to such content that its existence can be evidenced.

A consequence of such redundancy is that my analysis can only give a partial insight into and will focus necessarily on the negative areas of discourse (i.e. disagreement), where latent content and background common knowledge is inadequate to support aligned marking decisions.

This means that there is still a methodological gap around how to realise the extent of TLs' and examiners' background common knowledge use, and how they integrate this knowledge into their on-going feedback interactions.

Another concern I have about my methodology is that the unfolding nature of research project, and in particular the insights that were gained towards the latter end of the analysis process, left some unresolved issues that I was unable to deal with. As I discussed earlier, the use of semi-structured interviews entails a trade-off between researcher dominance (i.e. imposing a focus on the discussion), and eliciting unanticipated, yet valid, views from the participants.

Following the quantitative data analyses I noticed that there were two findings that I could not explain through my theoretical framework. I was not able to anticipate the use of *Directive* and *Technical Information* in Economics feedback in particular, and this may need additional refinement of my theoretical framework or further data collection to probe this finding. As it stands, this is a gap in my analysis and represents a weakness in the way that I integrated my quantitative and my qualitative data analysis so to inform each other.

I outlined above how the analysis of *Time* was a key feature of my theoretical framework. I chose to divide my corpus into four equal segments to give me insights into change and development. This choice was pragmatic (allowing me to manageably construct comparisons in my dataset through the Monoconc MP 2.2 concordancing and corpus analysis software (Barlow, 2012)). I am unable to claim that the inferences made from these analyses would be

the same if I had subdivided the corpus in a different way, and so this is a feature that could require additional research attention.

Another limitation relates to the potential for my professional role to have influenced the participants (and the nature of the reflections that they chose to share) within the study. Although I took steps to guard against my professional researcher role being an influence on the participants (e.g. through the careful development of interview tools and distancing my affiliated links with the OCR awarding body), it is possible that the examiners in the study perceived a degree of social influence that I could not discount. There could also be concerns about the self-selective nature of the examiners who opted into the study. Although only four of the 34 examiners who were approached to participate in the research project declined, the reasons for their refusal are unknown and it is possible that the nature of their interactions would have differed from the cohort of examiners who participated in my study.

Finally, the opportunistic nature of the sampling design in my study meant that some examiner relationship factors could not be explained with confidence through my analysis (e.g. the influence of experience on interaction). These elements would need to be explored further in subsequent study. Sampling and selection choices are key elements of how qualitative studies demonstrate authenticity, and this has implications for the generalisability of their findings (LeCompte & Preissle, 1993). My choice to use established examiner team structures as the basis for my data collection foregrounded the importance of capturing naturally occurring data at the expense of constructing a sample of participants who demonstrated a spread of characteristics that could allow me to systematically test for variances in feedback discourse (e.g. across examiners with different levels of experience or familiarity).

This is an area where my work could be extended further. The limitations of the study design meant that it was not possible to fully separate the interactions between TL-examiner familiarity and examiner experience when considering feedback construction. Additional work could be carried out to separate these dimensions through analysing the feedback content of examiner teams using a purposive sampling approach. This would allow statistical methods to be used to compare the analytical outcomes for the different examiner groups. These outcomes would also enable any current training guidance to be further refined and to potentially target it towards a particular examiner type.

My discussion of the methodological contributions and the limitations of my study now lead to my consideration of the implications and next steps.

7.2.3 Implications and Next Steps

The broad aim of the project was to help to improve the reliability of examination script marking through identifying the features of feedback communication that make it effective in terms of aligning the perspectives of TLs and examiners (and attaining the resolution of alternative examiner perspectives).

My IAF is a useful tool for articulating the components of feedback discourse in remote hierarchic professional environments. In the context of the study of examiner feedback this is important as this type of discourse is a hitherto hidden area of practice. My analyses can be used to highlight the features of practice that can structure senior examiner training in this area. The IAF and my analyses highlight how feedback includes the transaction of content (i.e. information giving around the location and the reason for mark credit, or technical information) and the use of references to things beyond the immediate context (i.e. Bridging). My analysis also shows that feedback was generally dealing with negative information (disagreement) and that there was more feedback around trouble sources (e.g. where complex marking decisions were being made, and where shared common ground between the examiners was limited).

As well as anticipating where examiners can expect the main focus of feedback work to be (i.e. around levels-based mark schemes and new or unfamiliar examiners), the IAF and my analyses also imply that examiner training needs to emphasise the importance of ensuring that examiner interactions foster and develop common ground. The reason for this relates to the measure of effectiveness that pertains to my theoretical framework, which is the ability of feedback to attain resolution.

Trouble sources are generally marked by disagreement in feedback discourse. As a means of ensuring that the transactional elements of feedback are successful (and resolution attained), the establishment and development of common ground requires a focus on the interactional elements of discourse. These interactional elements involve the use of articulation work, where the examiners employ negative politeness strategies to ensure that others' professional status is respected. This articulation work facilitates effective transaction through enabling on-going discourse that employs reference to shared resources (e.g. common texts and

conceptual content), the use of clear and specific language, and the incorporation of elaborated rationales.

The insights that I have developed as part of this study have already started to inform feedback practices. Drawing on the content of my academic output (Johnson, 2014a, 2015, 2016c, 2016d) and my assessment industry-related publications (Johnson, 2016a, 2016e), the insights from my analyses have fed into the development of a Feedback Training Module for examiner managers in the *Cambridge Assessment International Education* group and a new training package for OCR examination mark reviewers. I have also presented my analyses to Ofqual, the body involved in influencing examinations policies at a national level (Johnson, 2016b), and thereby contributing to debates around the public understanding of the stability of marks and the limits of examiner agreement is certain situations. A central element of this input has been to reiterate that feedback-giving has complex dimensions and that it has an impact on the development of a professional community.

Finally, it would be useful to extend the methodology to explore its potential for analysing remote video feedback. Although giving feedback to examiners by video messaging is not a practice that is currently available to examiners in the study context, it is possible that technological advancements will make this option available in the future. Such a development would allow examiners to access some of the paralinguistic cues that are currently unavailable to them, and would represent a shift towards a more rich feedback interaction. In advance of this, it would be useful to carry out a critical review of ASCDA methodology to consider whether it could be refined to include visual data.

8. References

AbuSeileek, A. F., & Qatawneh, K. (2013). Effects of synchronous and asynchronous computer-mediated communication (CMC) oral conversations on English language learners' discourse functions. *Computers & Education*, *62*, 181–190. https://doi.org/10.1016/j.compedu.2012.10.013

Ackerman, D. S., & Gross, B. L. (2010). Instructor Feedback: How Much Do Students Really Want? *Journal of Marketing Education*, *32*(2), 172–181. https://doi.org/10.1177/0273475309360159

Adcroft, A. (2011). The mythology of feedback. *Higher Education Research & Development*, *30*(4), 405–419. https://doi.org/10.1080/07294360.2010.526096

Ahmed, M., Sevdalis, N., Paige, J., Paragi-Gururaja, R., Nestel, D., & Arora, S. (2012). Identifying best practice guidelines for debriefing in surgery: a tri-continental study. *The American Journal of Surgery*, *203*(4), 523–529. https://doi.org/10.1016/j.amjsurg.2011.09.024

Ajjawi, R., & Boud, D. (2017). Researching feedback dialogue: an interactional analysis approach. *Assessment & Evaluation in Higher Education*, *42*(2), 252–265. https://doi.org/10.1080/02602938.2015.1102863

Ajjawi, R., Schofield, S., McAleer, S., & Walker, D. (2013). Assessment and feedback dialogue in online distance learning. *Medical Education*, *47*(5), 527–528. https://doi.org/10.1111/medu.12158

Alexander, R. (2010). Speaking but not listening? Accountable talk in an unaccountable context. *Literacy*, *44*(3), 103–111. https://doi.org/10.1111/j.1741-4369.2010.00562.x

Alvesson, M. (2011). *Interpreting Interviews*. London: SAGE.

Archer, J. C. (2010). State of the science in health professional education: effective feedback. *Medical Education*, *44*(1), 101–108. https://doi.org/10.1111/j.1365-2923.2009.03546.x

Asmuß, B. (2008). Performance Appraisal Interviews Preference Organization in Assessment Sequences. *Journal of Business Communication*, *45*(4), 408–429. https://doi.org/10.1177/0021943608319382

Atkinson, R., & Flint, J. (2001). Accessing Hidden and Hard-to-Reach Populations: Snowball Research Strategies. *Social Research UPDATE*, (33), 1–4.

Ayaß, R. (2015). Doing data: The status of transcripts in Conversation Analysis. *Discourse Studies*, *17*(5), 505–528. https://doi.org/10.1177/1461445615590717

Baker, P. (2004). Querying Keywords Questions of Difference, Frequency, and Sense in Keywords Analysis. *Journal of English Linguistics*, *32*(4), 346–359. https://doi.org/10.1177/0075424204269894

Bakhtin, M. M. (1981). *The dialogic imagination: four essays*. (M. Holquist, Ed.) (Vol. no. 1). Austin: University of Texas Press.

Bamberg, M. (1997). Positioning between structure and performance. *Journal of Narrative & Life History*, *7*(1–4), 335–342.

Barlow, M. (2012). Monoconc Pro (Version 2.2). Houston, TX: Athelstan.

Barnes, B. R. (2010). The Hawthorne Effect in community trials in developing countries. *International Journal of Social Research Methodology*, *13*(4), 357–370. https://doi.org/10.1080/13645570903269096

Bartolomé, F. (1989). Nobody Trusts the Boss Completely - Now What? *Harvard Business Review*, *67*(2), 135–142.

Barton, E. E., & Wolery, M. (2007). Evaluation of E-mail Feedback on the Verbal Behaviors of Pre-service Teachers. *Journal of Early Intervention*, *30*(1), 55–72. https://doi.org/10.1177/105381510703000105

Basturkmen, H., East, M., & Bitchener, J. (2014). Supervisors' on-script feedback comments on drafts of dissertations: socialising students into the academic discourse community. *Teaching in Higher Education*, *19*(4), 432–445. https://doi.org/10.1080/13562517.2012.752728

Bearman, M., Molloy, E., Ajjawi, R., & Keating, J. (2013). 'Is there a Plan B?': clinical educators supporting underperforming students in practice settings. *Teaching in Higher Education*, *18*(5), 531–544. https://doi.org/10.1080/13562517.2012.752732

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2007). The analysis of negotiation of common ground in CSCL. *Learning and Instruction*, *17*(4), 427–435. https://doi.org/10.1016/j.learninstruc.2007.04.002

Beighton, C., Poma, S., & Leonard, V. (2014). Expansive learning in firearms training: the HE/professional learning interface. *Police Practice and Research: An International Journal*, 1–13. https://doi.org/10.1080/15614263.2014.951045

Bennett, R. E. (2007). What Does It Mean to Be a Nonprofit Educational Measurement Organization in the 21st Century? In *Proceedings of the 33rd Annual Meeting of the International Association for Educational Assessment, Mutual Dependence of National Estimation Systems and Education Standards*. Baku, Azerbaijan. Retrieved from https://www.ets.org/Media/Research/pdf/Nonprofit.pdf

Bernstein, B. (2000). *Pedagogy, Symbolic Control, and Identity: Theory, Research, Critique*. Oxford: Rowman & Littlefield.

Bhatia, V. K. (1993). *Analysing Genre: Language Use in Professional Settings*. London: Pearson Education.

Bietz, M. J. (2008). Effects of Communication Media on the Interpretation of Critical Feedback. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (pp. 467–476). New York, NY, USA: ACM. https://doi.org/10.1145/1460563.1460637

Billett, S. (1993). Authenticity and a culture of practice. *Australian and New Zealand Journal of Vocational Education Research*, *2*(1), 1–29.

Billig, M. (1999). Whose Terms? Whose Ordinariness? Rhetoric and Ideology in Conversation Analysis. *Discourse & Society*, *10*(4), 543–558. https://doi.org/10.1177/0957926599010004005

Black, B., Curcin, M., & Dhawan, V. (2010). Investigating seeding items used for monitoring on-line marking: factors affecting marker agreement with the gold standard marks. Presented at the Annual Conference of the International Association for Educational Assessment (IAEA), August 2010, Bangkok, Thailand.

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Blair, A., & McGinty, S. (2013). Feedback-dialogues: exploring the student perspective. *Assessment & Evaluation in Higher Education*, *38*(4), 466–476. https://doi.org/10.1080/02602938.2011.649244

Bloom, B. S. (1953) Thought Processes in Lectures and Discussion. *The Journal of General Education, 7,* 160-169.

BNC Consortium. (2003). *Design of BNC-baby*. Oxford: Oxford University Computing Services. Retrieved from http://www.natcorp.ox.ac.uk/archive/oldBabyDocs/baby-des.html

BNC Consortium. (2007). British National Corpus (Version 3). Oxford: Oxford University Computing Services. Retrieved from http://www.natcorp.ox.ac.uk/

Boje, D. M. (2001). *Narrative Methods for Organizational & Communication Research*. London: SAGE Publications Ltd.

Boland, R. J., & Tenkasi, R. V. (1995). Perspective Making and Perspective Taking in Communities of Knowing. *Organization Science*, *6*(4), 350–372. https://doi.org/10.1287/orsc.6.4.350

Boling, E. C., Hough, M., Krinsky, H., Saleem, H., & Stevens, M. (2012). Cutting the distance in distance education: Perspectives on what promotes positive, online learning experiences. *The Internet and Higher Education*, *15*(2), 118–126. https://doi.org/10.1016/j.iheduc.2011.11.006

Bosley, S., & Young, D. (2006). Online Learning Dialogues in Learning through Work. *Journal of Workplace Learning*, *18*(6), 355–366.

Boud, D., & Molloy, E. (Eds.). (2013a). *Feedback in Higher and Professional Education: Understanding it and doing it well*. Abingdon, Oxon: Routledge.

Boud, D., & Molloy, E. (2013b). What is the problem with feedback? In D. Boud & E. Molloy (Eds.), *Feedback in Higher and Professional Education: Understanding it and doing it well* (pp. 1–10). Abingdon, Oxon: Routledge.

Bowe, H., & Martin, K. (2007). *Communication Across Cultures: Mutual Understanding in a Global World*. Cambridge: Cambridge University Press.

Bower, M. (2008). Affordance analysis – matching learning tasks with learning technologies. *Educational Media International*, *45*(1), 3–15. https://doi.org/10.1080/09523980701847115

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development* (Vol. xvi). Thousand Oaks, CA: Sage Publications, Inc.

Bramley, T. (2008). Mark scheme features associated with different levels of marker agreement. Presented at the British Educational Research Association annual conference, Heriot-Watt University, Edinburgh.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2016). (Mis)conceptualising themes, thematic analysis, and other problems with Fugard and Potts' (2015) sample-size tool for thematic analysis. *International Journal of Social Research Methodology*, *19*(6), 739–743. https://doi.org/10.1080/13645579.2016.1195588

Bridge, P., & Appleyard, R. (2005). System failure: A comparison of electronic and paper-based assignment submission, marking, and feedback. *British Journal of Educational Technology*, *36*(4), 669–671. https://doi.org/10.1111/j.1467-8535.2005.00485.x

British Educational Research Association. (2012). Ethical Guidelines for Educational Research. BERA. Retrieved from www.bera.ac.uk/publications/Ethical Guidelines

Brown, A., & Dowling, P. (1998). *Doing Research, Reading Research: A Mode of Interrogation for Education*. London: The Falmer Press.

Brown, E., & Glover, C. (2006). Evaluating written feedback. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp. 81–91). London; New York: Routledge,.

Brown, G., & Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics*, *32*(10), 1439–1465. https://doi.org/10.1016/S0378-2166(99)00094-6

Burke, L., Marks-Maran, D. J., Ooms, A., Webb, M., & Cooper, D. (2009). Towards a pedagogy of work-based learning: perceptions of work-based learning in foundation degrees. *Journal of Vocational Education & Training*, *61*(1), 15–33. https://doi.org/10.1080/13636820902819917

Burr, V. (2003). *Social Constructionism* (2nd ed.). Hove, UK: Routledge.

Butler Shay, S. (2004). The Assessment of Complex Performance: A Socially Situated Interpretive Act. *Harvard Educational Review*, *74*(3), 307–329. https://doi.org/10.17763/haer.74.3.wq16l67103324520

Calderhead, J. (1981). Stimulated Recall: A Method for Research on Teaching. *British Journal of Educational Psychology*, *51*(2), 211–217. https://doi.org/10.1111/j.2044-8279.1981.tb02474.x

Cambridge Assessment. (2015). Information Security Policy. Cambridge Assessment.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1–47.

Carpenter, T. D., & Wisecarver, M. M. (2004). *Indentifying and Validating a Model of Interpersonal Performance Dimensions* (Technical Report No. 1144) (p. 52). Alexandria VA.: U.S. Army Research Institute for the Behavioral and Social Sciences.

Cassaday, H. J., Bloomfield, R. E., & Hayward, N. (2002). Relaxed conditions can provide memory cues in both undergraduates and primary school children. *The British Journal of Educational Psychology*, *72*(4), 531–547.

Cassell, J. (1980). Ethical Principles for Conducting Fieldwork. *American Anthropologist*, *82*(1), 28–41. https://doi.org/10.2307/676126

Chetwynd, F., & Dobbyn, C. (2011). Assessment, feedback and marking guides in distance education. *Open Learning: The Journal of Open, Distance and e-Learning*, *26*(1), 67–78. https://doi.org/10.1080/02680513.2011.538565

Chur-Hansen, A., & McLean, S. (2006). On being a Supervisor: The Importance of Feedback and how to give it. *Australasian Psychiatry*, *14*(1), 67–71. https://doi.org/10.1080/j.1440-1665.2006.02248.x

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Clark, H. H. (1992). *Arenas of Language Use*. Stanford, CA: University of Chicago Press.

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in Communication. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on Socially Shared Meaning* (pp. 127–149). Washington, DC: American Psychological Association.

Cohen, J. (1969). *Statistical power analysis for the behavioural sciences*. London: Academic Press.

Cohen, L., & Manion, L. (1994). *Research Methods in Education* (4th ed.). London; New York: Routledge.

Collins, P., Yao, X., & Borlongan, A. (2014). Relative clauses in Philippine English: a diachronic perspective. In L. Vandelanotte, K. Davidse, C. Gentens, & D. Kimps (Eds.), *Recent advances in Corpus Linguistics: Developing and Exploiting Corpora* (pp. 125–146). Amsterdam: Rodopi.

Coninx, N., Kreijns, K., & Jochems, W. (2013). The use of keywords for delivering immediate performance feedback on teacher competence development. *European*

Journal of Teacher Education, *36*(2), 164–182. https://doi.org/10.1080/02619768.2012.717613

Conole, G., & Dyke, M. (2004). What are the affordances of information and communication technologies? *ALT-J*, *12*(2), 113–124. https://doi.org/10.1080/0968776042000216183

Cook, C. A., O'Shea, A. M., Young, A. L., & Stedmon, A. W. (1999). Helpful Feedback? Deriving Guidelines for Advanced Training Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *43*(21), 1133–1137. https://doi.org/10.1177/154193129904302104

Cook, D. L. (1962). The Hawthorne Effect in Educational Research. *The Phi Delta Kappan*, *44*(3), 116–122.

Corbin, J., & Morse, J. M. (2003). The unstructured interactive interview: Issues of reciprocity and risks when dealing with sensitive topics. *Qualitative Inquiry*, *9*(3), 335.

Cornelius, C., & Boos, M. (2003). Enhancing Mutual Understanding in Synchronous Computer-Mediated Communication by Training Trade-Offs in Judgmental Tasks. *Communication Research*, *30*(2), 147–177. https://doi.org/10.1177/0093650202250874

Crabtree, A., Tolmie, P., & Rouncefield, M. (2013). "How Many Bloody Examples Do You Want?" Fieldwork and Generalisation. In O. W. Bertelsen, L. Ciolfi, M. A. Grasso, & G. A. Papadopoulos (Eds.), *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus* (pp. 1–20). London: Springer.

Cramton, C. D. (2001). The Mutual Knowledge Problem and Its Consequences for Dispersed Collaboration. *Organization Science*, *12*(3), 346–371. https://doi.org/10.1287/orsc.12.3.346.10098

Crisp, V., & Johnson, M. (2007). The use of annotations in examination marking: opening a window into markers' minds. *British Educational Research Journal*, *33*(6), 943–961. https://doi.org/10.1080/01411920701657066

Culpeper, J. (2001). *Language and Characterisation: People in Plays and Other Texts*. Abingdon, Oxon: Routledge.

Culpeper, J., & Haugh, M. (2014). *Pragmatics and the English Language*. Basingstoke: Palgrave Macmillan. Retrieved from http://https://he.palgrave.com/page/detail/pragmatics-and-the-english-language-jonathan-culpeper/?sf1=barcode&st1=9780230551732

Dannels, D. P. (2000). Learning to Be Professional Technical Classroom Discourse, Practice, and Professional Identity Construction. *Journal of Business and Technical Communication*, *14*(1), 5–37. https://doi.org/10.1177/105065190001400101

Dannels, D. P., & Martin, K. N. (2008). Critiquing Critiques A Genre Analysis of Feedback Across Novice to Expert Design Studios. *Journal of Business and Technical Communication*, *22*(2), 135–159. https://doi.org/10.1177/1050651907311923

Davis, B. D., & Foster, P. (2002). Performance Documentation. *Business Communication Quarterly*, *65*(2), 108–114. https://doi.org/10.1177/108056990206500212

de Haan, P., & van Esch, K. (2004). Towards an Instrument for the Assessment of the Development of Writing Skills. In U. Connor & T. A. Upton (Eds.), *Applied Corpus Linguistics: A Multidimensional Perspective* (pp. 267–279). Amsterdam: Rodopi.

Dearnley, C. A., Taylor, J. D., Laxton, J. C., Rinomhota, S., & Nkosana-Nyawata, I. (2013). The student experience of piloting multi-modal performance feedback tools in health and social care practice (work)-based settings. *Assessment & Evaluation in Higher Education*, *38*(4), 436–450. https://doi.org/10.1080/02602938.2011.645014

Dennen, V. P. (2004). Cognitive apprenticeship in educational practice: Research on scaffolding, modeling, mentoring, and coaching as instructional strategies. In D. H. Jonassen (Ed.), *Handbook of Research on Educational Communications and Technology* (2nd ed., pp. 813–828). Mahwah, NJ: Lawrence Erlbaum Associates.

Dennen, V. P., Aubteen Darabi, A., & Smith, L. J. (2007). Instructor–Learner Interaction in Online Courses: The relative perceived importance of particular instructor actions on performance and satisfaction. *Distance Education*, *28*(1), 65–79. https://doi.org/10.1080/01587910701305319

Dennis, A. R., Fuller, R. M., & Valacich, J. S. (2008). Media, Tasks, and Communication Processes: A Theory of Media Synchronicity. *MIS Q.*, *32*(3), 575–600.

Dennis, A. R., & Valacich, J. S. (1999). Rethinking media richness: towards a theory of media synchronicity. Presented at the 32nd Annual Hawaii International Conference on Systems Sciences, 5-8 January 1999, Washington, DC: IEEE Computer Society. https://doi.org/10.1109/HICSS.1999.772701

Denscombe, M. (2010). *The Good Research Guide* (4th ed.). Maidenhead: Open University Press.

Department of Basic Education. (2009). *National Examinations and Assessment: Report on the National Senior Certificate Examination Results (Part 2)* (p. 260). Pretoria, South Africa: Department of Basic Education. Retrieved from http://www.education.gov.za/LinkClick.aspx?fileticket=l3hlVk9sypk%3D&...

Derks, D., & Bakker, A. B. (2010). The Impact of E-mail Communication on Organizational Life. *Cyberpsychology : Journal of Psychosocial Research on Cyberspace*, *4*(1). Retrieved from http://www.cyberpsychology.eu/view.php?cisloclanku=2010052401&..

DeSanctis, G., & Monge, P. (1998). Communication Processes for Virtual Organizations. *Journal of Computer-Mediated Communication, 3*(4). https://doi.org/10.1111/j.1083-6101.1998.tb00083.x

Dibble, J. L., & Levine, T. R. (2010). Breaking Good and Bad News: Direction of the MUM Effect and Senders' Cognitive Representations of News Valence. *Communication Research, 37*(5), 703–722. https://doi.org/10.1177/0093650209356440

Dickey, M. H., Wasko, M. M., Chudoba, K. M., & Bennett Thatcher, J. (2006). Do You Know What I Know? A Shared Understandings Perspective on Text-Based Communication. *Journal of Computer-Mediated Communication, 12*(1), 66–87. https://doi.org/10.1111/j.1083-6101.2006.00315.x

Donnellon, A., Gray, B., & Bougon, M. G. (1986). Communication, Meaning, and Organized Action. *Administrative Science Quarterly, 31*(1), 43–55.

Dowden, T., Pittaway, S., Yost, H., & McCarthy, R. (2013). Students' perceptions of written feedback in teacher education: ideally feedback is a continuing two-way communication that encourages progress. *Assessment & Evaluation in Higher Education*, *38*(3), 349–362. https://doi.org/10.1080/02602938.2011.632676

Edwards, A. (2012). The role of common knowledge in achieving collaboration across practices. *Learning, Culture and Social Interaction, 1*(1), 22–32. https://doi.org/10.1016/j.lcsi.2012.03.003

Edwards, D., & Mercer, N. (1987). *Common Knowledge: The Development of Understanding in the Classroom*. London: Methuen.

Ellis, R. A., Goodyear, P., Prosser, M., & O'Hara, A. (2006). How and what university students learn through online and face-to-face discussion: conceptions, intentions and approaches. *Journal of Computer Assisted Learning*, *22*(4), 244–256. https://doi.org/10.1111/j.1365-2729.2006.00173.x

Engeström, Y. (1987). *Learning by Expanding: an activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsultit.

Engeström, Y. (2001). Expansive Learning at Work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, *14*(1), 133–156. https://doi.org/10.1080/13639080020028747

Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology*, *70*(1), 113–136. https://doi.org/10.1348/000709900158001

Espasa, A., & Meneses, J. (2010). Analysing feedback processes in an online teaching and learning environment: an exploratory study. *Higher Education*, *59*(3), 277–292. https://doi.org/10.1007/s10734-009-9247-4

Evans, C. (2013). Making Sense of Assessment Feedback in Higher Education. *Review of Educational Research*, *83*(1), 70–120. https://doi.org/10.3102/0034654312474350

Evans, G. T., & Butler, J. (1992). Expert models and feedback processes in developing competence in industrial trade areas. *Australian Journal of TAFE Research and Development*, *8*(1), 13–32.

Evans, J., & Benefield, P. (2001). Systematic Reviews of Educational Research: Does the medical model fit? *British Educational Research Journal*, *27*(5), 527–541. https://doi.org/10.1080/01411920120095717

Fereday, J., & Muir-Cochrane, E. (2008). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods*, *5*(1), 80–92.

Festinger, L. (1957). *A theory of cognitive conflict*. Stanford, CA: Stanford University Press.

Festinger, L., & Katz, D. (1953). *Research methods in the behavioral sciences*. New York: Holt, Rinehart and Winston.

Field, A., & Hole, G. (2003).*How to Design and Report Experiments*. London: Sage Publications Ltd.

Firmin, M. W. (2008). Themes. In L. M. Given (Ed.), *The Sage Encyclopedia of Qualitative Research Methods* (pp. 869–870). Thousand Oaks, CA: SAGE Publications, Inc.

Foddy, W. (1993). *Constructing Questions for Interviews and Questionnaires*. Cambridge: Cambridge University Press.

Fowler, H. W., & Fowler, F. G. (Eds.). (1974). *The Concise Oxford Dictionary* (5th ed.). Oxford: Oxford University Press.

Fox-Turnbull, W. (2009). Stimulated recall using autophotography-A method for investigating technology education (pp. 204–217). Presented at the PATT-22 Conference, Delft, the Netherlands.

Frankfort-Nachmias, C., & Nachmias, D. (1996). *Research Methods in the Social Sciences* (5th ed.). London: Arnold.

Fransen, M. L., & Hoeven, C. L. ter. (2013). Matching the Message The Role of Regulatory Fit in Negative Managerial Communication. *Communication Research*, *40*(6), 818–837. https://doi.org/10.1177/0093650211427140

Franzosi, R. (2004). Encyclopedia of Social Science Research Methods. In M. S. Lewis-Beck, A. Bryman, & T. Liao (Eds.), *Content Analysis* (Vols 1–3, pp. 187–191). Thousand Oaks, CA: SAGE Publications, Inc. Retrieved from http://knowledge.sagepub.com/view/socialscience/n166.xml

Fraser, B., & Nolen, W. (2009). The association of deference with linguistic form. *International Journal of the Sociology of Language*, *1981*(27), 93–110. https://doi.org/10.1515/ijsl.1981.27.93

Furnborough, C., & Truman, M. (2009). Adult beginner distance language learner perceptions and use of assignment feedback. *Distance Education*, *30*(3), 399–418. https://doi.org/10.1080/01587910903236544

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Garven, S., Wood, J. M., & Malpass, R. S. (2000). Allegations of wrongdoing: the effects of reinforcement on children's mundane and fantastic claims. *The Journal of Applied Psychology*, *85*(1), 38–49.

Garven, S., Wood, J. M., Malpass, R. S., & Shaw III, J. S. (1998). More than suggestion: The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology*, *83*(3), 347–359. https://doi.org/10.1037/0021-9010.83.3.347

Gass, S. M. (2001). Innovations in Second Language Research Methods. *Annual Review of Applied Linguistics*, *21*, 221–32.

Gass, S. M., & Mackey, A. (2000). *Stimulated Recall Methodology in Second Language Research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gasson, S. (2005). The Dynamics of Sensemaking, Knowledge, and Expertise in Collaborative, Boundary-Spanning Design. *Journal of Computer-Mediated Communication*, *10*(4). https://doi.org/10.1111/j.1083-6101.2005.tb00277.x

Geddes, D., & Linnehan, D. (1996). Exploring the dimensionality of positive and negative performance feedback. *Communication Quarterly*, *44*(3), 326–344. https://doi.org/10.1080/01463379609370021

Gee, J. P., & Green, J. L. (1998). Discourse Analysis, Learning, and Social Practice: A Methodological Study. *Review of Research in Education*, *23*, 119–169. https://doi.org/10.2307/1167289

Georgakopoulou, A. (2011). “On for drinkies?”: Email cues of participant alignments. *Language@Internet*, *8*(4). Retrieved from http://www.languageatinternet.org/articles/2011/Georgakopoulou

Gibson, W., & Brown, A. (2009). *Working with Qualitative Data*. London: SAGE Publications Ltd.

Gibson, W., Webb, H., & vom Lehn, D. (2011). Re-constituting social praxis: an ethnomethodological analysis of video data in optometry consultations. *International Journal of Social Research Methodology*, *14*(3), 207–218. https://doi.org/10.1080/13645579.2011.563618

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Paris, France: Editions de la Maison des Sciences de l’Homme.

Giordan, M. (2003). The role of IRF exchanges in the discursive dynamics of e-mail tutored interactions. *International Journal of Educational Research*, *39*(8), 817–827. https://doi.org/10.1016/j.ijer.2004.11.004

Goffman, E. (1955). On face-work: an analysis of ritual elements in social interaction. *Psychiatry: Journal for the Study of Interpersonal Processes*, *18*, 213–231.

Goffman, E. (1967). *Interaction Ritual: Essays in Face to Face Behavior*. Chicago: Aldine.

Golden, P. A., Beauclair, R., & Sussman, L. (1992). Factors affecting electronic mail use. *Computers in Human Behavior*, *8*(4), 297–311. https://doi.org/10.1016/0747-5632(92)90026-B

Gomm, R., Hammersley, M., & Foster, P. (2000). *Case study method: key issues, key texts*. Thousand Oaks, Calif.: SAGE.

Gould, J., & Day, P. (2013). Hearing you loud and clear: student perspectives of audio feedback in higher education. *Assessment & Evaluation in Higher Education*, *38*(5), 554–566. https://doi.org/10.1080/02602938.2012.660131

Greatorex, J., & Bell, J. F. (2008). What makes AS marking reliable? An experiment with some stages from the standardisation process. *Research Papers in Education*, *23*(3), 333–355. https://doi.org/10.1080/02671520701692593

Green, J., Franquiz, M., & Dixon, C. (1997). The Myth of the Objective Transcript: Transcribing as a Situated Act. *TESOL Quarterly*, *31*(1), 172–176. https://doi.org/10.2307/3587984

Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *British Medical Journal*, *331*(7524), 1064–1065. https://doi.org/10.1136/bmj.38636.593461.68

Gregory, J. B., & Levy, P. E. (2012). Employee feedback orientation: implications for effective coaching relationships. *Coaching: An International Journal of Theory, Research and Practice*, *5*(2), 86–99. https://doi.org/10.1080/17521882.2012.691888

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied Thematic Analysis*. Thousand Oaks, CA: SAGE Publications, Inc.

Guichon, N., Bétrancourt, M., & Prié, Y. (2012). Managing written and oral negative feedback in a synchronous online teaching situation. *Computer Assisted Language Learning*, *25*(2), 181–197. https://doi.org/10.1080/09588221.2011.636054

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English* (1st ed.). Harlow, UK: Longman.

Hamilton, L., & Corbett-Whittier, C. (2013). *Using Case Study in Education Research*. London: SAGE.

Hammersley, M., & Traianou, A. (2012). *Ethics in Qualitative Research: Controversies and Contexts*. London: SAGE Publications Ltd.

Hancock, J. T., & Dunham, P. J. (2001). Language Use in Computer-Mediated Communication: The Role of Coordination Devices. *Discourse Processes*, *31*(1), 91–110. https://doi.org/10.1207/S15326950dp3101_4

Harms, P. L., & Roebuck, D. B. (2010). Teaching the Art and Craft of Giving and Receiving Feedback. *Business Communication Quarterly*, *73*(4), 413–431. https://doi.org/10.1177/1080569910385565

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hatzipanagos, S., & Warburton, S. (2009). Feedback as dialogue: exploring the links between formative assessment and social software in distance learning. *Learning, Media and Technology*, *34*(1), 45–59. https://doi.org/10.1080/17439880902759919

Hebert, B. G., & Vorauer, J. D. (2003). Seeing through the screen: is evaluative feedback communicated more effectively in face-to-face or computer-mediated exchanges? *Computers in Human Behavior*, *19*(1), 25–38. https://doi.org/10.1016/S0747-5632(02)00031-6

Henderson, P., Ferguson-Smith, A. C., & Johnson, M. H. (2005). Developing essential professional skills: a framework for teaching and learning about feedback. *BMC Medical Education*, *5*(1), 1–6. https://doi.org/10.1186/1472-6920-5-11

Heritage, J. (2004). Conversation analysis and institutional talk: Analyzing data. In D. Silverman (Ed.), *Qualitative Analysis: Theory, Method and Practice* (Second, pp. 222–245). London: SAGE.

Herring, S. C. (1999). Interactional coherence in CMC. Presented at the 32nd Annual Hawaii International Conference on Systems Sciences, 1999. HICSS-32. https://doi.org/10.1109/HICSS.1999.772674

Higgins, R., Hartley, P., & Skelton, A. (2001). Getting the Message Across: The problem of communicating assessment feedback. *Teaching in Higher Education*, *6*(2), 269–274. https://doi.org/10.1080/13562510120045230

Hoey, E. M., & Kendrick, K. H. (2017). Conversation Analysis. In A. M. B. de Groot & P. Hagoort (Eds.), *Research Methods in Psycholinguistics: A Practical Guide.* Chichester: Wiley-Blackwell.

Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. London; New York: Routledge.

Holmes, J. (2001). *An Introduction to Sociolinguistics* (2nd ed.). Harlow, UK: Pearson Education.

Holmes, J., & Stubbe, M. (2003). *Power and Politeness in the Workplace: A Sociolinguistic Analysis of Talk at Work*. London: Longman.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley Pub. Co.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117. https://doi.org/10.1016/0010-0277(96)81418-1

Hruschka, D. J., Schwartz, D., St.John, D. C., Picone-Decaro, E., Jenkins, R. A., & Carey, J. W. (2004). Reliability in Coding Open-Ended Data: Lessons Learned from HIV Behavioral Research. *Field Methods*, *16*(3), 307–331. https://doi.org/10.1177/1525822X04266540

Hudson, P. (2014). Feedback consistencies and inconsistencies: eight mentors' observations on one preservice teacher's lesson. *European Journal of Teacher Education*, *37*(1), 63–73. https://doi.org/10.1080/02619768.2013.801075

Huffaker, D. (2010). Dimensions of Leadership and Social Influence in Online Communities. *Human Communication Research*, *36*(4), 593–617. https://doi.org/10.1111/j.1468-2958.2010.01390.x

Hutchins, S., McDermott, P., Carolan, T., Gronowski, M., Fisher, A., & DeMay, M. (2013). *Interpersonal Skills Summary Report* (Interim No. 2013-03) (p. 147). Fort Belvoir: US Army Research Institute for the Behavioral and Social Sciences. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a585788.pdf

Hyatt, D. F. (2005). 'Yes, a very good point!': a critical genre analysis of a corpus of feedback commentaries on Master of Education assignments. *Teaching in Higher Education*, *10*(3), 339–353. https://doi.org/10.1080/13562510500122222

Hyland, F. (2001). Providing Effective Support: Investigating feedback to distance language learners. *Open Learning: The Journal of Open, Distance and e-Learning*, *16*(3), 233–247. https://doi.org/10.1080/02680510120084959

Ilyenkov, E. V. (2008). *Dialectical Logic; Essays on its History and Theory*. Delhi: Aakar Books.

Ingram, J., & Elliott, V. (2016). A critical analysis of the role of wait time in classroom interactions and the effects on student and teacher interactional behaviours. *Cambridge Journal of Education*, *46*(1), 37–53. https://doi.org/10.1080/0305764X.2015.1009365

Jarvenpaa, S. L., & Leidner, D. E. (1998). Communication and Trust in Global Virtual Teams. *Journal of Computer-Mediated Communication*, *3*(4). https://doi.org/10.1111/j.1083-6101.1998.tb00080.x

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13–23). Philadelphia: John Benjamins Publishing.

Johlke, M. C., & Duhan, D. F. (2000). Supervisor communication practices and service employee job outcomes. *Journal of Service Research*, *3*(2), 154–165.

Johnson, M. (2007). Does the anticipation of a merit grade motivate vocational test-takers? *Research in Post-Compulsory Education*, *12*(2), 159–179. https://doi.org/10.1080/13596740701387445

Johnson, M. (2008). Exploring assessor consistency in a Health and Social Care qualification using a sociocultural perspective. *Journal of Vocational Education & Training*, *60*(2), 173–187. https://doi.org/10.1080/13636820802042446

Johnson, M. (2013, October). *Examiner feedback and learning - What are the characteristics of effective remote feedback in a hierarchic, professional context* (PhD Registration Proposal Report). University of Cambridge, Cambridge.

Johnson, M. (2014a). A case study of inter-examiner feedback from a UK context: Mixing research methods to gain insights into situated learning interactions. *Formation et Pratiques d'enseignement En Questions*, *17*, 67–88.

Johnson, M. (2014b). Insights into Contextualised Learning: How Do Professional Examiners Construct Shared Understanding through Feedback? *E-Learning and Digital Media*, *11*(4), 363–378. https://doi.org/10.2304/elea.2014.11.4.363

Johnson, M. (2015). Articulation work: Insights into examiners' expertise from their remote feedback interactions. *Communication & Language at Work*, *1*(4), 28–52.

Johnson, M. (2016a). All in good time: Influences on team leaders' feedback communication choices when giving feedback to examiners. *Research Matters: A Cambridge Assessment Publication*, *21*, 28–33.

Johnson, M. (2016b). Examining examiner work: Evidencing complex practices. Presented at the Ofqual Symposium: How consistent can marking be?, London.

Johnson, M. (2016c). Feedback effectiveness in professional learning contexts. *Review of Education*, *4*(2), 195–229. https://doi.org/10.1002/rev3.3061

Johnson, M. (2016d). Reading between the Lines: Exploring Methods for Analysing Professional Examiner Feedback Discourse. *International Journal of Research & Method in Education*, *40*(5), 456-470. https://doi.org/10.1080/1743727X.2016.1166484

Johnson, M. (2016e). The challenges of researching digital technology use: examples from an assessment context. *International Journal of E-Assessment*, *1*(2), 1–10.

Johnson, M., & Black, B. (2012a). Feedback as scaffolding: Senior examiner monitoring processes and their effects on examiner marking. *Research in Post-Compulsory Education*, *17*(4), 391–407. https://doi.org/10.1080/13596748.2012.738965

Johnson, M., & Black, B. (2012b). What's Going On? Analysing Visual Data to Understand Context-Based Decision-Making Processes. *International Journal of Research & Method in Education*, *35*(3), 243–250. https://doi.org/10.1080/1743727X.2012.713222

Johnson, M., & Shaw, S. (2012). Interpreting examiners' annotations on examination papers: a sociocultural analysis. *Irish Educational Studies*, *31*(4), 467–485. https://doi.org/10.1080/03323315.2012.739828

Johnston, B. (2004). Summative assessment of portfolios: an examination of different approaches to agreement over outcomes. *Studies in Higher Education*, *29*(3), 395–412. https://doi.org/10.1080/0307507041000168264 6

Johri, A. (2012). From a distance: Impression formation and impression accuracy among geographically distributed coworkers. *Computers in Human Behavior*, *28*(6), 1997–2006. https://doi.org/10.1016/j.chb.2012.06.038

Jordan, L. (2012). Video for peer feedback and reflection: embedding mainstream engagement into learning and teaching practice. *ALT-J: Research in Learning Technology*, *20*(:sup1), 16–25. https://doi.org/10.3402/rlt.v20i0.19192

Joyce, M. (2013). Picking the Best Intercoder Reliability Statistic for your Digital Activism Content Analysis. Retrieved 2 February 2015, from http://digital-activism.org/2013/05/picking-the-best-intercoder-reliability-statistic-for-your-digital-activism-content-analysis/

Kerssen-Griep, J., & Witt, P. L. (2015). Instructional Feedback III: How Do Instructor Facework Tactics and Immediacy Cues Interact to Predict Student Perceptions of Being Mentored? *Communication Education*, *64*(1), 1–24. https://doi.org/10.1080/03634523.2014.978797

Kim, D., Park, Y., Yoon, M., & Jo, I.-H. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. *The Internet and Higher Education*, *30*, 30–43. https://doi.org/10.1016/j.iheduc.2016.03.002

King, D., McGugan, S., & Bunyan, N. (2008). Does it make a difference? Replacing text with audio feedback. *Practice and Evidence of Scholarship of Teaching and Learning in Higher Education*, *3*(2), 145–163.

King, P. E. (2016). When Do Students Benefit From Performance Feedback? A Test of Feedback Intervention Theory in Speaking Improvement. *Communication Quarterly*, *64*(1), 1–15. https://doi.org/10.1080/01463373.2015.1078827

Kirk, D. S., & Sellen, A. (2010). On Human Remains: Values and Practice in the Home Archiving of Cherished Objects. *ACM Trans. Comput.-Hum. Interact.*, *17*(3), 10:1–10:43. https://doi.org/10.1145/1806923.1806924

Kirschner, P. A., & Lai, K. (2007). Online communities of practice in education. *Technology, Pedagogy and Education*, *16*(2), 127–131. https://doi.org/10.1080/14759390701406737

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Knight, S., & Littleton, K. (2013). Discourse, computation and context – sociocultural DCLA revisited. Presented at the 1st International Workshop on Discourse-Centric Learning Analytics 2013, Leuven, Belgium. Retrieved from http://oro.open.ac.uk/36640/

Knoch, U., Read, J., & von Randow, J. (2007). Re-Training Writing Raters Online: How Does It Compare with Face-to-Face Training? *Assessing Writing*, *12*(1), 26–43.

Korzenny, F. (1978). A Theory of Electronic Propinquity Mediated Communication in Organizations. *Communication Research*, *5*(1), 3–24. https://doi.org/10.1177/009365027800500101

Koschmann, T., Kuutti, K., & Hickman, L. (1998). The Concept of Breakdown in Heidegger, Leont'ev, and Dewey and Its Implications for Education. *Mind, Culture, and Activity*, *5*(1), 25–41. https://doi.org/10.1207/s15327884mca0501_3

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: SAGE.

Kruger, J., Epley, N., Parker, J., & Ng, Z.-W. (2005). Egocentrism over e-mail: can we communicate as well as we think? *Journal of Personality and Social Psychology*, *89*(6), 925–936. https://doi.org/10.1037/0022-3514.89.6.925

Krussel, L., Edwards, B., & Springer, G. T. (2004). The Teacher's Discourse Moves: A Framework for Analyzing Discourse in Mathematics Classrooms. *School Science and Mathematics*, *104*(7), 307–312. https://doi.org/10.1111/j.1949-8594.2004.tb18249.x

Kupferberg, I. (2013). Introduction: Institutional computer-mediated troubles talk. *Language@Internet*, *10*(2). Retrieved from http://www.languageatinternet.org/articles/2013/kupferberg

Kvale, S. (1996). *InterViews: An Introduction to Qualitative Research Interviewing*. London: SAGE Publications.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lakoff, G. (1973). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, *2*(4), 458–508. https://doi.org/10.1007/BF00262952

Lakoff, R. (1973). Language and Woman's Place. *Language in Society*, *2*(1), 45–80.

Lakoff, R. (1979). Stylistic Strategies Within a Grammar of Style. *Annals of the New York Academy of Sciences*, *327*(1), 53–78. https://doi.org/10.1111/j.1749-6632.1979.tb17753.x

Lambertz, K. (2011). Back-channelling: the use of yeah and mm to portray engaged listenership. *Griffiths Working Papers in Pragmatics and Intercultural Communication*, *4*(1/2), 11–18.

Laming, D. (2004). *Human Judgment: The Eye of the Beholder*. London: Thomson Learning.

Landsberger, H. A. (1958). *Hawthorne revisited: Management and the worker: its critics, and developments in human relations in industry*. Ithaca, NY: Cornell University.

Lapadat, J. C. (2007). Discourse Devices Used to Establish Community, Increase Coherence, and Negotiate Agreement in an Online University Course. *The Journal of Distance Education / Revue de l'Éducation à Distance*, *21*(3), 59–92.

Lapadat, J. C., & Lindsay, A. C. (1999). Transcription in Research and Practice: From Standardization of Technique to Interpretive Positionings. *Qualitative Inquiry*, *5*(1), 64–86. https://doi.org/10.1177/107780049900500104

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Lawrence, P. R. (1992). The Challenge of Problem-Oriented Research. *Journal of Management Inquiry*, *1*(2), 139–142. https://doi.org/10.1177/105649269212007

LeCompte, M., & Preissle, J. (1993). *Ethnography and qualitative design in educational research* (2nd ed.). London: Academic Press.

Lee, A. M., Landkin, D. K., & Carter, J. A. (1992). Student thoughts during tennis instruction. *Journal of Teaching in Physical Education*, *11*, 256–267.

Lefstein, A., Snell, J., & Israeli, M. (2015). From moves to sequences: expanding the unit of analysis in the study of classroom discourse. *British Educational Research Journal*, *41*(5), 866–885. https://doi.org/10.1002/berj.3164

Li, J., & De Luca, R. (2014). Review of assessment feedback. *Studies in Higher Education*, *39*(2), 378–393. https://doi.org/10.1080/03075079.2012.709494

Liberman, N., Sagristano, M. D., & Trope, Y. (2002). The effect of temporal distance on level of mental construal. *Journal of Experimental Social Psychology*, *38*(6), 523–534. https://doi.org/10.1016/S0022-1031(02)00535-8

Limerick, B., Burgess-Limerick, T., & Grace, M. (1996). The politics of interviewing: power relations and accepting the gift. *International Journal of Qualitative Studies in Education*, *9*(4), 449–460. https://doi.org/10.1080/0951839960090406

Lindberg, M. A., Keiffer, J., & Thomas, S. W. (2000). Eyewitness Testimony for Physical Abuse as a Function of Personal Experience, Development, and Focus of Study. *Journal of Applied Developmental Psychology*, *21*(5), 555–591. https://doi.org/10.1016/S0193-3973(00)00054-X

Linderbaum, B. A., & Levy, P. E. (2010). The Development and Validation of the Feedback Orientation Scale (FOS). *Journal of Management*, *36*(6), 1372–1405. https://doi.org/10.1177/0149206310373145

Lingard, L., Schryer, C., Garwood, K., & Spafford, M. (2003). 'Talking the talk': school and workplace genre tension in clerkship case presentations. *Medical Education*, *37*(7), 612–620. https://doi.org/10.1046/j.1365-2923.2003.01553.x

Littleton, K., & Mercer, N. (2013). *Interthinking: putting talk to work*. Abingdon, Oxon: Routledge.

Lombard, M., Snyder-Duch, J., & Campanella Bracken, C. (2010). Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects. Temple University, Philadelphia, PA, USA.

Lotman, Y. M. (1988). Text within a Text. *Soviet Psychology*, *26*(3), 32–51. https://doi.org/10.2753/RPO1061-0405260332

Lu, C.-J., & Shulman, S. W. (2008). Rigor and flexibility in computer-based qualitative research: Introducing the Coding Analysis Toolkit. *International Journal of Multiple Research Approaches*, *2*(1), 105–117. https://doi.org/10.5172/mra.455.2.1.105

Lurie, N. H., & Swaminathan, J. M. (2009). Is timely information always better? The effect of feedback frequency on decision making. *Organizational Behavior and Human Decision Processes*, *108*(2), 315–329. https://doi.org/10.1016/j.obhdp.2008.05.005

Lyle, J. (2003). Stimulated recall: a report on its use in naturalistic research. *British Educational Research Journal*, *29*(6), 861–878.

Mancilla, R. L., Polat, N., & Akcay, A. O. (2017). An Investigation of Native and Nonnative English Speakers' Levels of Written Syntactic Complexity in Asynchronous Online Discussions. *Applied Linguistics*, *38*(1), 112–134. https://doi.org/10.1093/applin/amv012

Martin, D., & Rouncefield, M. (2003). Making the organization come alive: talking through and about the technology in remote banking. *Human-Computer Interaction*, *18*(1), 111–148. https://doi.org/10.1207/S15327051HCI1812_5

Martins, L. L., Gilson, L. L., & Maynard, M. T. (2004). Virtual Teams: What Do We Know and Where Do We Go From Here? *Journal of Management*, *30*(6), 805–835. https://doi.org/10.1016/j.jm.2004.05.002

Massey, A. J., & Raikes, N. (2006). Item level examiner agreement. Presented at the British Educational Research Association annual conference, University of Warwick, UK.

Mathieson, K. (2012). Exploring Student Perceptions of Audiovisual Feedback via Screencasting in Online Courses. *American Journal of Distance Education*, *26*(3), 143–156.

Mathisen, P. (2012). Video feedback in higher education - a contribution to improving the quality of written feedback. *Nordic Journal of Digital Literacy*, *7*(2), 97–113.

May, T. (2011). *Social Research: Issues, Methods and Research* (4th ed.). Maidenhead: McGraw-Hill International.

May, T. (2013). Identifying the characteristics of written formative feedback used by assessors in work-based qualifications. *Journal of Vocational Education & Training*, *65*(1), 18–32. https://doi.org/10.1080/13636820.2012.727855

Maybin, J. (2001). Language, Stuggle and Voice: The Bakhtin/Vološinov Writings. In M. Wetherell, S. Taylor, & S. Yates (Eds.), *Discourse theory and practice: a reader* (pp. 64–71). London: SAGE.

Meadows, M., & Billington, L. (2005). *A review of the literature on marking reliability*. London: National Assessment Agency. Retrieved from https://cerp.aqa.org.uk/sites/default/files/pdf_upload/CERP_RP_MM_01052005.pdf

Mehan, H. (1979). *Learning Lessons : Social Organization in the Classroom*. Cambridge, Mass: Harvard University Press.

Mengis, J., & Eppler, M. J. (2008). Understanding and Managing Conversations from a Knowledge Perspective: An Analysis of the Roles and Rules of Face-to-face Conversations in Organizations. *Organization Studies*, *29*(10), 1287–1313. https://doi.org/10.1177/0170840607086553

Mercer, N. (2000). *Words and Minds: How We Use Language to Think Together and Get Things Done*. London: Routledge.

Mercer, N. (2002). Developing Dialogues. In G. Wells & G. Claxton (Eds.), *Learning for Life in the 21st Century: Sociological Perspectives of the Future* (pp. 141–153). Oxford: Blackwell Publishers Ltd.

Mercer, N. (2004). Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, *1*(2), 137–168. https://doi.org/10.1558/japl.v1i2.137

Mercer, N. (2008a). Talk and the Development of Reasoning and Understanding. *Human Development*, *51*(1), 90–100. https://doi.org/10.1159/000113158

Mercer, N. (2008b). The Seeds of Time: Why Classroom Dialogue Needs a Temporal Analysis. *Journal of the Learning Sciences*, *17*(1), 33–59. https://doi.org/10.1080/10508400701793182

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Oxford: Routledge.

Mercer, N., Littleton, K., & Wegerif, R. (2004). Methods for Studying the Processes of Interaction and Collaborative Activity in Computer-Based Educational Activities. *Technology, Pedagogy and Education*, *13*(2), 195–212. https://doi.org/10.1080/14759390400200180

Mercer, N., & Wegerif, R. (1999). Is 'exploratory talk' productive talk? In K. Littleton & P. Light (Eds.), *Learning With Computers: Analysing Productive Interaction* (pp. 79–101). London: Routledge.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's Talk and the Development of Reasoning in the Classroom. *British Educational Research Journal*, *25*(1), 95–111. https://doi.org/10.1080/0141192990250107

Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education: Revised and Expanded from Case Study Research in Education*. San Francisco, CA: Jossey-Bass Publishers.

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative Discourse Idealized and Realized: Accountable Talk in the Classroom and in Civic Life. *Studies in Philosophy and Education*, *27*(4), 283–297. https://doi.org/10.1007/s11217-007-9071-1

Mirador, J. F. (2000). A Move Analysis of Written Feedback in Higher Education. *RELC Journal*, *31*(1), 45–60. https://doi.org/10.1177/003368820003100103

Moore, M. G. (1993). Theory of transactional distance. In D. Keegan (Ed.), *Theoretical Principles of Distance Education* (pp. 22–38). London: Routledge.

Moore, M. G. (2007). The Theory of Transactional Distance. In M. G. Moore (Ed.), *The Handbook of Distance Education* (2nd ed., pp. 89–108). Mahwah, NJ: Lawrence Erlbaum Associates.

Morand, D. A. (1996). Dominance, Deference, and Egalitarianism in Organizational Interaction: A Sociolinguistic Analysis of Power and Politeness. *Organization Science*, *7*(5), 544–556.

Morand, D. A. (2000). Language and power: an empirical analysis of linguistic strategies used in superior–subordinate communication. *Journal of Organizational Behavior*, *21*(3), 235–248. https://doi.org/10.1002/(SICI)1099-1379(200005)21:3<235::AID-JOB9>3.0.CO;2-N

Morgan, N., & Pritchard, A. (2005). On souvenirs and metonymy: Narratives of memory, metaphor and materiality. *Tourist Studies*, *5*(1), 29–53. https://doi.org/10.1177/1468797605062714

Morgan, S. J., & Symon, G. (2002). Computer-Mediated Communication and Remote Management: Integration or Isolation? *Social Science Computer Review*, *20*(3), 302–311. https://doi.org/10.1177/089443930202000307

Mory, E. H. (2004). Feedback Research Revisited. In *Handbook of Research on Educational Communications and Technology* (2nd ed., pp. 745–783). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from http://www.aect.org/edtech/29.pdf

Münzer, S., & Holmer, T. (2009). Bridging the Gap Between Media Synchronicity and Task Performance: Effects of Media Characteristics on Process Variables and Task Performance Indicators in an Information Pooling Task. *Communication Research*, *36*(1), 76–103. https://doi.org/10.1177/0093650208326464

Murphy, E. A., & Rodriguez, M. A. (2008). Revisiting Transactional Distance Theory in a Context of Web-Based High-School Distance Education. *The Journal of Distance Education / Revue de l'Éducation à Distance*, *22*(2). Retrieved from http://www.jofde.ca/index.php/jde/article/view/38

Murphy, P. (2010). Web-based collaborative reading exercises for learners in remote locations: the effects of computer-mediated feedback and interaction via computer-mediated communication. *ReCALL*, *22*(02), 112–134. https://doi.org/10.1017/S0958344010000030

Myers, G. (2004). *Matters of Opinion: Talking About Public Issues*. Cambridge: Cambridge University Press.

Nestel, D., Bello, F., & Kneebone, R. (2013). Feedback in clinical procedural skills simulations. In D. Boud & E. Molloy (Eds.), *Higher and Professional Education. Understanding It and Doing It Well* (pp. 140–157). Abingdon, Oxon: Routledge.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: SAGE.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Norris, S. P. (1990) Effect of eliciting verbal reports of thinking on critical thinking test performance. *Journal of Educational Measurement 27,* 41-58.

Ochs, E. (1979). Transcription as theory. In E. Ochs & B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 43–72). New York: Academic Press.

OCR. (2012). Instructions for On-Screen Marking – Marking from Home - 2013. Oxford Cambridge and RSA Examinations.

OCR. (2013). Marking and Grading. Oxford Cambridge and RSA Examinations.

Ofqual. (2011). *GCSE, GCE, Principal Learning and Project Code of Practice*. Coventry: Ofqual. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/371268/2011-05-27-code-of-practice.pdf

Ofqual. (2013). *Review of Quality of Marking in Exams in A levels, GCSEs and Other Academic Qualifications: Interim Report* (No. Ofqual/13/5287). Coventry: Office of Qualifications and Examinations Regulation. Retrieved from http://dera.ioe.ac.uk/17772/1/2013-06-07-review-of-quality-of-marking-in-exams-in-a-levels-gcses-and-other-academic-qualifications-interim-report.pdf

Örnberg Berglund, T. (2009). Disrupted Turn Adjacency and Coherence Maintenance in Instant Messaging Conversations. *Language@Internet*, *6*(2). Retrieved from http://www.languageatinternet.org/articles/2009/2106

Owens, J., Hardcastle, L. A., & Richardson, B. (2009). Learning From a Distance: The Experience of Remote Students. *The Journal of Distance Education / Revue de l'Éducation à Distance*, *23*(3), 53–74.

Panahi, P., Birjandi, P., & Azabdaftari, B. (2013). Toward a sociocultural approach to feedback provision in L2 writing classrooms: the alignment of dynamic assessment and teacher error feedback. *Language Testing in Asia*, *3*(1), 1–10. https://doi.org/10.1186/2229-0443-3-13

Paul, A., Gilbert, K., & Remedio, L. (2013). Socio-cultural considerations in feedback. In D. Boud & E. Molloy (Eds.), *Feedback in Higher and Professional Education: Understanding it and doing it well* (pp. 72–89). Abingdon, Oxon: Routledge.

Paulus, T. M. (1999). The Effect of Peer and Teacher Feedback on Student Writing. *Journal of Second Language Writing*, *8*(3), 265–289. https://doi.org/10.1016/S1060-3743(99)80117-9

Pearson. (2013). Assessment and information. Retrieved 25 September 2013, from http://www.pearsonassessments.com/pai/ea/research/researchandresources.htm

Pinkett, R. D. (2000). Bridging the digital divide: Sociocultural constructionism and an asset-based approach to community technology and community building. Presented at the

81st Annual Meeting of the American Educational Research Association (AERA), New Orleans.

Polanyi, M. (1969). *Personal Knowledge* (3rd ed.). London: Routledge & Kegan Paul.

Pomerantz, A. (1975). *Second Assessments: A Study of Some Features of Agreements/Disagreements*. University of California, Irvine.

Pomerantz, A. (1985). Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes. In M. Atkinson & H. Heritage (Eds.), *Structures of Social Action* (pp. 57–101). Cambridge: Cambridge University Press. Retrieved from http://dx.doi.org/10.1017/CBO9780511665868.008

Pomerantz, A., & Heritage, J. (2012). Preference. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 210–228). Chichester: John Wiley & Sons, Ltd. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/9781118325001.ch11/summary

Prins, F. J., Sluijsmans, D. M. A., & Kirschner, P. A. (2006). Feedback for General Practitioners in Training: Quality, Styles, and Preferences. *Advances in Health Sciences Education*, *11*(3), 289–303. https://doi.org/10.1007/s10459-005-3250-z

Pryor, E., & Woodward-Kron, R. (2014). International medical graduate doctor to doctor telephone communication: A genre perspective. *English for Specific Purposes*, *35*, 41–53. https://doi.org/10.1016/j.esp.2013.12.001

Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, *34*(1), 1–20. https://doi.org/10.1080/03054980701476386

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, *28*(1), 4–13. https://doi.org/10.1002/bs.3830280103

Raz, A. E., & Fadlon, J. (2006). Managerial Culture, Workplace Culture and Situated Curricula in Organizational Learning. *Organization Studies*, *27*(2), 165–182. https://doi.org/10.1177/0170840605056399

Recio Saucedo, A., Whitelock, D. M., Gilbert, L. H., Hatzipanagos, S., Gillary, P., Watt, S., & Zhang, P. (2013). Improving a tutor's feedback assessment tool: transforming Open Mentor following two recent deployments. *International Journal of E-Assessment*, *3*(1). Retrieved from http://journals.sfu.ca/ijea/index.php

Ren, Y., Kraut, R., & Kiesler, S. (2007). Applying Common Identity and Bond Theory to Design of Online Communities. *Organization Studies*, *28*(3), 377–408. https://doi.org/10.1177/0170840607076007

Renz, A. D., Watts, C. A., & Conrad, D. A. (2013). In M. A. Keating, A. McDermott, & K. Montgomery (Eds.), *Patient-Centred Health Care: Achieving Co-ordination, Communication and Innovation* (pp. 186–196). New York, NY: Palgrave Macmillan.

RM Education. (2014). RM Assessor. Retrieved 3 February 2014, from http://rmresults.co.uk/content/rm-assessor

Rojas-Drummond, S., Torreblanca, O., Pedraza, H., Vélez, M., & Guzmán, K. (2013). 'Dialogic scaffolding': Enhancing learning and understanding in collaborative contexts. *Learning, Culture and Social Interaction*, *2*(1), 11–21. https://doi.org/10.1016/j.lcsi.2012.12.003

Ros i Solé, C., & Truman, M. (2005). Feedback in Distance Learning Programmes in Languages: Attitudes to linguistic faults and implications for the learning process. *Distance Education*, *26*(3), 299–323. https://doi.org/10.1080/01587910500291389

Rovai, A. P. (2000). Building and sustaining community in asynchronous learning networks. *The Internet and Higher Education*, *3*(4), 285–297. https://doi.org/10.1016/S1096-7516(01)00037-9

Sacks, H. (1963). Sociological Description. *Berkeley Journal of Sociology*, *8*, 1–16.

Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button & J. R. Lee (Eds.), *Talk and Social Organisation* (pp. 54–69). Clevedon: Multilingual Matters.

Sacks, H. (1992). *Lectures on Conversation*. (G. Jefferson & E. A. Schegloff, Eds.) (Volumes I and II edition). Oxford: Wiley-Blackwell.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/BF00117714

Salager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes*, *13*(2), 149–170. https://doi.org/10.1016/0889-4906(94)90013-2

Salerno, S. M., O'Malley, P. G., Pangaro, L. N., Wheeler, G. A., Moores, L. K., & Jackson, J. L. (2002). Faculty Development Seminars Based on the One-Minute Preceptor Improve Feedback in the Ambulatory Setting. *Journal of General Internal Medicine*, *17*(10), 779–787. https://doi.org/10.1046/j.1525-1497.2002.11233.x

Santamaría-García, C. (2017). Emotional and Educational Consequences of (Im)politeness in Teacher–Student Interaction at Higher Education. *Corpus Pragmatics*, 1–23. https://doi.org/10.1007/s41701-017-0010-2

Sawchuk, P. H., & Stetsenko, A. (2008). Sociological Understandings of Conduct for a Noncanonical Activity Theory: Exploring Intersections and Complementarities. *Mind, Culture, and Activity*, *15*(4), 339–360. https://doi.org/10.1080/10749030802391427

Schegloff, E. A. (1992). Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. *American Journal of Sociology*, *97*(5), 1295–1345.

Schegloff, E. A. (1997). Whose Text? Whose Context? *Discourse & Society*, *8*(2), 165–187. https://doi.org/10.1177/0957926597008002002

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language*, *53*(2), 361–382.

Schegloff, E. A., & Sacks, H. (1973). Opening up Closings. *Semiotica*, *8*(4), 289–327.

Schmidt, K. (1994). Cooperative Work and its Articulation: Requirements for Computer Support. *Le Travail Humain*, *57*(4), 345–366.

Schmidt, K. (2011). *Cooperative Work and Coordinative Practices: Contributions to the Conceptual Foundations of Computer-Supported Cooperative Work (CSCW)*. London: Springer-Verlag.

Schön, D. A. (1983). *The Reflective Practitioner How Professionals Think in Action*. London: Temple Smith.

Schwarz, B. B., Kolikant, Y. B.-D., & Mishenkina, M. (2012). 'Co-alienation' mediated by common representations in synchronous e-discussions. *Learning, Culture and Social Interaction*, *1*(3–4), 216–231. https://doi.org/10.1016/j.lcsi.2012.09.002

Seedhouse, P. (2004). *The interactional architecture of the language classroom: a conversation analysis perspective.* Malden, MA: Blackwell.

Segoe, B. (2013). Feedback measures as a strategy for student support in open distance learning: Can it come to the aid of weary student teachers? *Mediterranean Journal of Social Sciences*, *4*(14), 705–710. https://doi.org/10.5901/mjss.2013.v4n14p705

Severinson Eklundh, K. (2010). To Quote or Not to Quote: Setting the Context for Computer-Mediated Dialogues. *Language@Internet*, *7*(5). Retrieved from http://www.languageatinternet.org/articles/2010/2665

Shalem, Y., & Slonimsky, L. (2010). Seeing Epistemic Order: Construction and Transmission of Evaluative Criteria. *British Journal of Sociology of Education*, *31*(6), 755–778.

Sias, P. M., Gallagher, E. B., Kopaneva, I., & Pedersen, H. (2012). Maintaining Workplace Friendships Perceived Politeness and Predictors of Maintenance Tactic Choice. *Communication Research, 39*(2), 239–268. https://doi.org/10.1177/0093650210396869

Silverman, D. (2004). *Qualitative Research: Theory, Method and Practice*. London: SAGE.

Sinclair, J., & Coulthard, M. (1975). *Towards an Analysis of Discourse: The English used by Teachers and Pupils*. Oxford: Oxford University Press.

Smith, P. J. (2003). Workplace Learning and Flexible Delivery. *Review of Educational Research, 73*(1), 53–88. https://doi.org/10.3102/00346543073001053

Smits, M., Boon, J., Sluijsmans, D., & van Gog, T. (2008). Content and timing of feedback in a web-based learning environment: effects on learning as a function of prior knowledge. *Interactive Learning Environments*, *16*(2), 183–193. https://doi.org/10.1080/10494820701365952

Social Research Association. (2003). *Ethical Guidelines*. London: Social Research Association. Retrieved from http://the-sra.org.uk/wp-content/uploads/ethics03.pdf

Spafford, M. M., Schryer, C. F., Mian, M., & Lingard, L. (2006). Look Who's Talking Teaching and Learning Using the Genre of Medical Case Presentations. *Journal of Business and Technical Communication, 20*(2), 121–158. https://doi.org/10.1177/1050651905284396

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge: Cambridge University Press.

Stannard, R. (2008). A New Direction in Feedback. *Humanising Language Teaching, 10*(6). Retrieved from http://www.hltmag.co.uk/dec08/mart04.htm

Stein, D. S., Wanstreet, C. E., Slagle, P., Trinko, L. A., & Lutz, M. (2013). From 'hello' to higher-order thinking: The effect of coaching and feedback on online chats. *The Internet and Higher Education*, *16*, 78–84. https://doi.org/10.1016/j.iheduc.2012.03.001

Stevenson, J. (2001). Vocational knowledge and its specification. *Journal of Vocational Education and Training*, *53*(4), 647–662. https://doi.org/10.1080/13636820100200187

Stivers, T. (2015). Coding Social Interaction: A Heretical Approach in Conversation Analysis? *Research on Language and Social Interaction*, *48*(1), 1–19. https://doi.org/10.1080/08351813.2015.993837

Stracke, E., & Kumar, V. (2010). Feedback and self-regulated learning: insights from supervisors' and PhD examiners' reports. *Reflective Practice*, *11*(1), 19–32. https://doi.org/10.1080/14623940903525140

Strauss, A. (1985). Work and the Division of Labor. *The Sociological Quarterly*, *26*(1), 1–19.

Suárez-Orozco, C., & Hernández, M. G. (2012). Immigrant Family Separations. In C. T. G. Coll (Ed.), *The Impact of Immigration on Children's Development* (pp. 122–148). Basel, Switzerland: Karger Medical and Scientific Publishers.

Sussman, S. W., & Sproull, L. (1999). Straight Talk: Delivering Bad News through Electronic Communication. *Information Systems Research*, *10*(2), 150–166. https://doi.org/10.1287/isre.10.2.150

Swales, J. (2014). Genre and Discourse Community. In J. Angermuller, D. Maingueneau, & R. Wodak (Eds.), *The Discourse Studies Reader: Main currents in theory and analysis* (pp. 305–316). Amsterdam: John Benjamins Publishing Company.

Swann, J. (2010). Transcribing spoken interaction. In S. Hunston & D. Oakey (Eds.), *Introducing Applied Linguistics: Concepts and Skills* (pp. 163–176). London: Routledge.

Tang, J., & Harrison, C. (2011). Investigating university tutor perceptions of assessment feedback: three types of tutor beliefs. *Assessment & Evaluation in Higher Education*, *36*(5), 583–604. https://doi.org/10.1080/02602931003632340

Taylor, J. D., Dearnley, C. A., Laxton, J. C., Nkosana-Nyawata, I., & Rinomhota, S. (2012). The Perceptions of Health and Social Care Students of using Mobile 360 Degree Performance Feedback Tools in Practice Placement Settings: *International Journal of Mobile and Blended Learning*, *4*(1), 18–30. https://doi.org/10.4018/jmbl.2012010102

Taylor, S. S. (2011). "I Really Don't Know What He Meant by That": How Well Do Engineering Students Understand Teachers' Comments on Their Writing? *Technical Communication Quarterly*, *20*(2), 139–166. https://doi.org/10.1080/10572252.2011.548762

Thomas, D. R. (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*, *27*(2), 237–246. https://doi.org/10.1177/1098214005283748

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, *39*, 212–222. https://doi.org/10.2307/1415413

Thurlings, M., Vermeulen, M., Kreijns, K., Bastiaens, T., & Stijnen, S. (2012). Development of the Teacher Feedback Observation Scheme: evaluating the quality of feedback in peer groups. *Journal of Education for Teaching*, *38*(2), 193–208. https://doi.org/10.1080/02607476.2012.656444

Tjeerdsma, B. L. (1997). A comparison of teacher and student perspectives of tasks and feedback. *Journal of Teaching in Physical Education*, *16*(4), 388–400.

Trees, A. R., Kerssen-Griep, J., & Hess, J. A. (2009). Earning Influence by Communicating Respect: Facework's Contributions to Effective Instructional Feedback. *Communication Education*, *58*(3), 397–416. https://doi.org/10.1080/03634520802613419

Trevino, L. K., Lengel, R. H., & Daft, R. L. (1987). Media Symbolism, Media Richness, and Media Choice in Organizations A Symbolic Interactionist Perspective. *Communication Research*, *14*(5), 553–574. https://doi.org/10.1177/009365087014005006

Tseng, F.-C., & Kuo, F.-Y. (2014). A study of social participation and knowledge sharing in the teachers' online professional community of practice. *Computers & Education*, *72*, 37–47. https://doi.org/10.1016/j.compedu.2013.10.005

Tuck, J. (2012). Feedback-giving as social practice: teachers' perspectives on feedback as institutional requirement, work and dialogue. *Teaching in Higher Education*, *17*(2), 209–221. https://doi.org/10.1080/13562517.2011.611870

Upton, T. A., & Cohen, M. A. (2009). An approach to corpus-based discourse analysis: The move analysis as example. *Discourse Studies*, *11*(5), 585–605. https://doi.org/10.1177/1461445609341006

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, *15*(3), 398–405. https://doi.org/10.1111/nhs.12048

van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes A Meta-Analysis. *Review of Educational Research*, *85*(4), 475–511. https://doi.org/10.3102/0034654314564881

Veloski, J., Boex, J. R., Grasberger, M. J., Evans, A., & Wolfson, D. B. (2006). Systematic review of the literature on assessment, feedback and physicians' clinical performance*: BEME Guide No. 7. *Medical Teacher*, *28*(2), 117–128. https://doi.org/10.1080/01421590600622665

VERBI Software – Consult – Sozialforschung GmbH. (2013). MAXqda, software for qualitative data analysis (Version 11). Marburg, Berlin.

Vološinov, V. N. (1973). *Marxism and the Philosophy of Language*. New York, NY, USA: Seminar Press Inc.

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: a case study. *The Internet and Higher Education*, *6*(1), 77–90. https://doi.org/10.1016/S1096-7516(02)00164-1

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, Mass: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language*. Massachusetts: MIT press.

Waldvogel, J. (2007). Greetings and Closings in Workplace Email. *Journal of Computer-Mediated Communication*, *12*(2), 456–477. https://doi.org/10.1111/j.1083-6101.2007.00333.x

Walvoord, A. A. G., Redden, E. R., Elliott, L. R., & Coovert, M. D. (2008). Empowering followers in virtual teams: Guiding principles from theory and practice. *Computers in Human Behavior*, *24*(5), 1884–1906. https://doi.org/10.1016/j.chb.2008.02.006

Wang, T., & Li, L. Y. (2011). 'Tell me what to do' vs. 'guide me through it': Feedback experiences of international doctoral students. *Active Learning in Higher Education*, *12*(2), 101–112. https://doi.org/10.1177/1469787411402438

Watts, S. A. (2007). Evaluative Feedback: Perspectives on Media Effects. *Journal of Computer-Mediated Communication*, *12*(2), 384–411. https://doi.org/10.1111/j.1083-6101.2007.00330.x

Weber, K., & Glynn, M. A. (2006). Making Sense with Institutions: Context, Thought and Action in Karl Weick's Theory. *Organization Studies*, *27*(11), 1639–1660. https://doi.org/10.1177/0170840606068343

Wegerif, R. (1997). Children's talk and computer software: a response to Fisher. In R. Wegerif & P. Scrimshaw (Eds.), *Computers and Talk in the Primary Classroom* (pp. 99–112). Clevedon: Multilingual Matters.

Wegerif, R. (2008). Dialogic or dialectic? The significance of ontological assumptions in research on educational dialogue. *British Educational Research Journal*, *34*(3), 347–361. https://doi.org/10.1080/01411920701532228

Wertsch, J. V., & Kazak, S. (2011). Saying More than You Know in Instructional Settings. In T. Koschmann (Ed.), *Theories of Learning and Studies of Instructional Practice* (Vol. 1, pp. 153–166). New York: Springer.

Wetherell, M. (1998). Positioning and Interpretative Repertoires: Conversation Analysis and Post-Structuralism in Dialogue. *Discourse & Society*, *9*(3), 387–412. https://doi.org/10.1177/0957926598009003005

Whitaker, B. G., Dahling, J. J., & Levy, P. E. (2007). The Development of a Feedback Environment and Role Clarity Model of Job Performance. *Journal of Management*, *33*(4), 570–591. https://doi.org/10.1177/0149206306297581

Whitelock, D. M., Gilbert, L., & Gale, V. (2011). Technology Enhanced Assessment and Feedback: How is evidence-based literature informing practice? Presented at the 2011 International Computer Assisted Assessment (CAA) Conference, Research into e-

Assessment, Southampton. Retrieved from http://caaconference.co.uk/wp-content/uploads/WhitelockB-CAA2011.pdf

Whitelock, D. M., Watt, S., Raw, Y., & Moreale, E. (2003). Analysing tutor feedback to students: First steps towards constructing an electronic monitoring system. *ALT-J: Research in Learning Technology*, *11*(3), 31–42. https://doi.org/10.1080/0968776030110304

Whitford, T., & Moss, S. A. (2009). Transformational Leadership in Distributed Work Groups: The Moderating Role of Follower Regulatory Focus and Goal Orientation. *Communication Research*, *36*(6), 810–837. https://doi.org/10.1177/0093650209346800

Whittaker, S. (2003). Things to talk about when talking about things. *Human-Computer Interaction*, *18*(1), 149–170. https://doi.org/10.1207/S15327051HCI1812_6

Wickström, G., & Bendix, T. (2000). The 'Hawthorne effect' - what did the original Hawthorne studies actually show? *Scandinavian Journal of Work, Environment & Health*, *26*(4), 363–367.

Wilcox, S., & Trudel, P. (1998). Constructing the coaching principles and beliefs of a youth ice hockey coach. *AVANTE*, *4*(3), 39–66.

Wiliam, D. (1993). Validity, dependability and reliability in National Curriculum assessment. *Curriculum Journal*, *4*(3), 335–350. https://doi.org/10.1080/0958517930040303

Wilson, C., & Williams, E. (1977). Watergate Words A Naturalistic Study of Media and Communication. *Communication Research*, *4*(2), 169–178. https://doi.org/10.1177/009365027700400203

Wilson, J. M., O'Leary, M. B., Metiu, A., & Jett, Q. R. (2008). Perceived Proximity in Virtual Work: Explaining the Paradox of Far-but-Close. *Organization Studies*, *29*(7), 979–1002. https://doi.org/10.1177/0170840607083105

Winter, E. C., & McGhie-Richmond, D. (2005). Using computer conferencing and case studies to enable collaboration between expert and novice teachers. *Journal of Computer Assisted Learning*, *21*(2), 118–129. https://doi.org/10.1111/j.1365-2729.2005.00119.x

Wiredu, G. O. (2007). Distance and contradictory motives in distributed workplace learning for a new profession. *Journal of Education and Work*, *20*(4), 357–377. https://doi.org/10.1080/13639080701464491

Wittgenstein, L. (2009). *Philosophical Investigations*. (P. M. S. Hacker & J. Schulte, Eds.) (4th Edition edition). Chichester: Wiley-Blackwell.

Yelland, C. (2011). A genre and move analysis of written feedback in higher education. *Language and Literature*, *20*(3), 218–235. https://doi.org/10.1177/0963947011413563

Yin, R. K. (1981). The Case Study as a Serious Research Strategy. *Science Communication*, *3*(1), 97–114. https://doi.org/10.1177/107554708100300106

Yin, R. K. (2013). *Case Study Research: Design and Methods* (5th ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Yinger, R. J. (1986). Examining thought in action: A theoretical and methodological critique of research on interactive teaching. *Teaching and Teacher Education*, *2*(3), 263–282. https://doi.org/10.1016/S0742-051X(86)80007-5

Ypsilandis, G. S. (2002). Feedback in Distance Education. *Computer Assisted Language Learning*, *15*(2), 167–181. https://doi.org/10.1076/call.15.2.167.8191

Zachariadis, M., Scott, S. V., & Barrett, M. I. (2010). *Exploring critical realism as the theoretical foundation of mixed-method research: evidence from the economies for IS innovations* (Monograph). Retrieved from http://www.jbs.cam.ac.uk/research/working_papers/index.html

Zemel, A., & Koschmann, T. (2011). Pursuing a Question: Reinitiating IRE Sequences as a Method of Instruction. *Journal of Pragmatics*, *43*(2), 475–488. https://doi.org/10.1016/j.pragma.2010.08.022

Appendix A: Glossary of Terms

AO	Assessment Organisation	*an organisation that designs, develops, delivers and awards the recognition of learning outcomes (knowledge, skills and/or competences) for an individual following an assessment. The term can be used synonymously with 'Awarding Body' or 'Exam Board'.*
	Assistant Examiner	*see 'Examiners'.*
AB	Awarding Body	*see 'Assessment Organisation'.*
	Cambridge Assessment	*a department of the University of Cambridge and the parent body of three examination boards; Oxford, Cambridge and RSA Examinations (OCR), University of Cambridge International Examinations (CIE), and Cambridge English Language Assessment.*
	Exam (Examination) Board	*see 'Assessment Organisation'.*
	Examiner	*examiners (sometimes termed 'Assistant Examiners') are responsible for marking candidates' work in accordance with the agreed mark scheme and marking procedures. Examiners are often current or retired teachers with experience of delivering the subject being examined.*
GCE	General Certificate of Education	*GCE courses are usually studied over a two-year period and are widely recognised in England, Wales and Northern Ireland as being the standard entry qualification for assessing the suitability of applicants for academic courses in UK Universities. GCE students are usually around 18 years old.*

	Item	*a ‘question’; any part of a test for which a mark or marks are given (a multiple choice question is an item, as is a gap to be filled in a sentence).*
	Levels-based Mark Scheme	*the total mark awarded is determined by matching the quality, content, and combination of elements of the candidate responses to the appropriate level of response, each of which is associated with a band of one or more marks. Markers have to apply a principle of best fit.*
	Mark Scheme	*a document laying out the specifications for the marking of a test.*
	Marking Panel	*a group of examiners who are appointed to mark an examination unit or component. For exams with a large entry population, the whole panel is subdivided into teams.*
	Objective Mark Scheme	*this is where an exact answer is a single requirement for marking. This is a binary (yes/no) decision for the examiner.*
Ofqual	Office of Qualifications and Examinations Regulation	*a non-ministerial government department with jurisdiction in England for regulating qualifications. Among other things the organisation is responsible for making sure that assessments and exams show what a student has achieved and that people have confidence in the regulated qualifications.*
OCR	Oxford, Cambridge and RSA Examinations	*an examination board that sets examinations and awards qualifications (including GCE A-levels). It is one of England, Wales and Northern Ireland’s five main examination boards.*
	Points-based Mark Scheme	*the total mark awarded is an accumulation of single marks awarded independently, through markers*

		locating the relevant elements in a (largely) prescriptive mark scheme.
PE	Principal Examiner	*the principal examiner for a qualification unit or component is responsible for the setting of the question paper and the standardising of its marking. They oversee the marking quality of Team Leaders (TLs).*
	scoris® assessor	*a web-based e-marking system used at OCR. There is a facility in the software for the examiners to electronically message each other. This is the most common method for communicating feedback.*
	Seed Script/Item	*seeds are the main way in which exam boards monitor the accuracy of examiners' marking. These are scripts/items that have already been marked and given a true score at pre-standardisation or at the pre-standardisation phase. Seeds appear at random in an examiner's marking allocation, without the examiner knowing that it is a seed, and must be marked by the examiner within tolerance. Seeds typically appear at a rate of about 5 per cent, which means that out of 20 scripts/items marked by an examiner, one is likely to be a seed.*
SSU	Standardisation Set Up Meeting	*a meeting where senior examiners (the qualification Principal Examiner and Team Leaders meet to select scripts that will be used to monitor the marking quality of examiners. The definitive mark for these scripts is decided through discussion at this meeting. This meeting also finalises the final/definitive mark scheme.*
TL	Team Leader	*senior examiners who oversee a group of up to around six subordinate examiners. TLs will*

normally have examined for at least two years in the same or a related subject. They attend the SSU meeting and input to the refinement of the final/definitive mark scheme. TLs report to a Principal Examiner (PE).

Appendix B: Journal Search Source List and Academic Domains

	Journal	Academic domain[18]				
		E	T	C	O	A
1	ALT-J: Research in Learning Technology	■	■			
2	Assessment & Evaluation in Higher Education	■				■
3	Assessment in Education: Principles, Policy & Practice	■				■
4	British Educational Research Journal	■				
5	British Journal of Educational Studies	■				
6	British Journal of Educational Technology	■	■			
7	Business and Professional Communication Quarterly	■	■			
8	Communication Research			■		
9	Communication Research Reports			■		
10	Computers & Education	■	■			
11	Computers in Human Behavior		■			
12	Cyberpsychology : journal of psychosocial research on cyberspace		■			
13	Distance Education	■	■			
14	Electronic Journal of e-Learning	■	■			
15	European Journal of Teacher Education	■				
16	E-learning and Digital Media*	■	■			
17	Human Communication Research			■		
18	Human-Computer Interaction		■			
19	Interactive Learning Environments	■	■			
20	International Journal of Educational Research	■				
21	International Journal of Human-Computer Studies		■			
22	Journal of Asynchronous Learning Networks	■	■			
23	Journal of Business and Technical Communication			■	■	
24	Journal of Computer Assisted Learning	■	■			
25	Journal of Computer-Mediated Communication		■	■		
26	Journal of Education and Work	■			■	
27	Journal of Higher Education	■				
28	Journal of Management				■	
29	Journal of Organizational Behavior				■	
30	Journal of Research on Technology in Education*	■	■			
31	Journal of Teaching and Learning in Higher Education*	■				
32	Journal of Vocational Education & Training	■			■	
33	Journal of Workplace Learning	■			■	
34	Language and Education	■		■		
35	Language@Internet		■	■		
36	Learning and Instruction	■				
37	Learning, Culture and Social Interaction	■				
38	Learning, Media and Technology	■	■			
39	Open Learning: The Journal of Open, Distance and e-Learning	■	■			
40	Organization Studies				■	
41	Organizational Behavior and Human Decision Processes	■			■	
42	Oxford Review of Education	■				
43	Research Papers in Education	■				
44	Review of Educational Research	■				
45	Scandinavian Journal of Educational Research	■				
46	Studies in Higher Education	■				
47	Teaching and Teacher Education	■				
48	Teaching in Higher Education	■				
49	technology, Pedagogy and Education	■	■			
50	The Internet and Higher Education	■	■			
51	The Journal of Distance Education / Revue de l'Éducation à Distance	■	■			
52	Western Journal of Communication			■		
53	Written Communication			■		

*Journal database non-searchable – journal removed from the search

[18] E: Education; T: Technology; C: Communication; O: Organisation; A: Assessment

Appendix C: Face-to-Face Meeting Observation Schedules [Pilot Study]

This schedule records the time taken to discuss each item, as well as gathering instances of in situ 'talk types'. These data will then be used to refine a schedule that might potentially gather qualitative indicators of Team Leader [TL] talk types. The meeting has 3 sessions:

Session 1 (modelling); Session 2 (group marking); Session 3 (individual marking). *During 'quiet' marking periods – collate notes on the types of talk/questions raised.*

Analysis

1. Time per item (collate average)
2. Examples of types of TL talk – questions raised
3. Examples of types of examiner-examiner talk
4. Examples of 'common knowledge' talk
5. Comparison of collected talk types, participation index figures with anticipated 'talk characteristics' and 'question forms' (validation and improvement)

Session 1 2 3

(a) Item/Time: (b) Talk Types (D. Edwards & Mercer, 1989, p. 6):

1. Offering new information
2. Referencing past experience
3. Requests for information
4. Checks on validity of interpretation of information
5. Other…?

Item	Time (start)	Time (finish)	NOTES: (TL → EX; EX → TL; EX → EX?)

Appendix D: Pilot Coding System [Pilot Study]

Primary code	Sub code	Notes	Example
1 Structure *Features of discourse that help to initially establish the direction of the communication*	Foreground TL	Preparing recipient for next part of the message	TL: I've got four points on this, none of which is earth shattering.
	Orientation: focus TL	TL drawing attention to common reference point	TL: So first of all, 2B.
	Orientation: focus EX	Examiner drawing attention to common reference point	EX: Okay I've got 3B looking at me

Primary code	Sub code	Notes	Example
2 Information *Where information is communicated*	?: Request TL	TL question: request for information	TL: Tell me when you are already on 3B.
	?: Request EX	Examiner question: request for information	EX: Are we looking at the diagram?
	1	Offering new information	TL: You gave it 46 and we gave it 46.
	+1	Reinforcement of new point/nascent ritual knowledge?	TL: By saying, "At point Y in the diagram, more is being spent on the restoration of houses." That's not good enough for the reference to the data mark. The reference to the data mark there must have reference to the axis on which the opportunity cost is shown. EX: Right, thank you. TL: So it's got to be, if they have coordinates going across to X and Y and they called it, sort of, you know, X1 and Y1.
	Principle/Perspective TL	TL gives principled knowledge: explanation	TL: My rule with my own students is, "Never, ever use the word money unless it's in an essay about monetary policy or macroeconomics and increasing the supply of money."
	No Perspective given		TL: We gave two comment marks.
	TL standard	Insight into TLs standards/signs of 'quality'	TL: Now, in point two, income, this is actually a standard normal, good answer.
	Principle/Perspective EX	Examiner gives principled knowledge: explanation	EX: Sure, I was giving that third mark, the second mark rather for the opportunity cost aspect, then obviously with which I shouldn't have been.
	Back Ref TL (asit-MS)	TL reference to past shared information: not co-present - e.g. MS reference	TL: They said, "The opportunity cost is Y1 - X1, that would have been fine," but don't, please don't give that 'reference to the diagram' point unless it's a reference to an opportunity cost shown on an axis.
	Back Ref TL (asit-cont/lang)	Content/concept knowledge - *as understood by the senior examiners*; 'familiar'/everyday language use	TL: 2A, this is a trap. "Opportunity cost means to choose the next best alternative forgone."
	Bridging (sit-	TL bringing together past and present information	TL: Now this bit of the conversation we had

	asit) TL	– contextualisation; e.g. Reference to performance ***and*** MS wording	before you'd arrived at the meeting so I'll just quickly through what I've said.
	Back Ref TL (shared)	Back Reference: past shared information/metadiscursive	TL: Because it's jumbled up money and houses, hence money is coming in again. EX: Okay. TL: Therefore, and then you've done the same thing with B that you did in the earlier scripts.
	Back Ref EX (asit-MS)	Examiner reference to past shared information: not co-present - e.g. MS reference	EX: It [the MS] does actually. It says, "Subsidies effectively reduced the cost, level three, brand one."
	Back Ref EX (shared)	Examiner reference to shared situation/object	EX: Okay. I do feel more confident with that [the MS concept] than I did at the meeting.
	Bridging (sit-asit) EX	Examiner bringing together past and present information – contextualisation; e.g. Reference to performance ***and*** MS wording	EX: I'd just put 'seen' on that is because I was counting the second determinant had been written in part one, saying that, "furthermore, more people are living alone," but with a second reason.
	Rhetorical elicitation	Question asked to elicit/steer information	TL: Yes. Why was it wrong to give it level three for the diagram?
	Senior Examiner Knowledge	Reference to information discussed at TL meeting	TL: We [senior examiners] voted on this I think and there was a majority of one for giving it a second mark.
	Marking convention	Conventions of distributed/large scale marking	TL: Please note that every page, including additional sheets, needs a "Seen" stamp.
	General marking information	General information about overall marks	TL: Well, I mean, in general the marking - your marking was very good. It was very clear what you were doing. It was very systematic.
	Self dialogue	Talking/thinking aloud	TL: Well, actually, what do we give it? Well, what you could actually do is, I mean, I don't think to be honest when we're at standardisation, I think actually we did look at it, does it say anything more?

Primary code	Sub code	Notes	Example
3 Disagreement / Misalignment *Where the focus of the discussion is around disagreement*	Dispreferred	Dispreferred Response	TL: You are absolutely right to give three to the diagram. EX: Yes. TL: But in the explanation they've got to represent the opportunity cost on the axes. EX: Right…
	Disagreement TL	Misalignment discussed by TL	TL: Therefore if the axes are invalid, you can't get a mark for referring to them in the explanation.
	Disagreement EX	Misalignment discussed by examiner	TL: You didn't like the second point but we did it give it one… EX: I didn't think it was another country, okay.
	Implied incorrect (TL)	TL implies that the examiner is not correct (without stating it)	TL: We gave two comment marks.
	Moderating disagreement	TL recognising ambiguity/complexity of decision making process; hints at the inability to capture elements of 'definite' outcomes in a MS/across multiple examiners	TL: So purely on a sort of, you know, logical pinicketiness, we didn't think that was valid because they have collated the statistics so it can't have been that difficult. But that is English rather than economics, really, and question six you were spot on.

Primary code	Sub code	Notes	Example
4 Style *The use of linguistic terms to situate the participants with respect to each other*	Distancing TL	TL moving 'self' from corporate interpretation; seeing/acknowledging the perspective of the examiner; recognising ambiguity in judgement making process	TL: We made that L3, band one. Actually we gave it ten, but looking at again, it shouldn't possibly have got more than nine really.
	Social TL	Knowledge link to wider community (non-exam); link to wider 'community understanding of knowledge' - ecological validity; e.g. how concepts taught	TL: But we [senior examiners] talked about this for a long time, actually because they put pound signs on the axis, the opportunity cost should be measured in terms of the number of houses given up, new houses.
	Authority reference (TL)	Reference to 'we': authority; potential allusion to power/community; ecological validity involved?	TL: You gave it 46 and we gave it 46.
	Authority reference (EX)	Reference to 'we'; acquiring collegiate/community language?	EX: Yes, so therefore we've got some level four stuff coming in.
	Social Common Ground	Shared knowledge - not task specific; Reference to background knowledge of each other	TL: Well, thank you. Thank you for struggling up to the meeting with the long distance and all the things. How is your husband?

Primary code	Sub code	Notes	Example
5 Verification *Checking the received understanding of a communicated message*	Acknowledgement EX	Examiner indication of agreement with previous point; has 'flow' function - encourages other participant to continue on same path of discourse	TL: So we didn't give it the fourth mark on the axes. EX: Right.
	Acknowledgement TL	Includes: letting the examiner know if there perspective or marking is correct; repetition of previous point made by the examiner	TL: So that is absolutely right to put a tick exactly where you've put it is spot on.
	Point acceptance TL	Agreement acknowledgement	
	Point acceptance EX	Point clearly taken/understood - nascent shift in perspective; powerful indicator of movement...?	TL: …Are you happy with that? EX: Yes. So that would have just been three marks? Plus one for the -
	Validity check TL	TL checks that their interpretation is correct	
	Validity check EX	Examiner checks that their interpretation is correct	EX: Okay, so that should be a four.
	?: Verify understanding TL	TL questions to elicit information on understanding	TL: Are you happy on that one?
	?: Verify understanding EX	Examiner questions to elicit information on understanding	EX: So on that first one, then they wouldn't get - ?

Primary code	Sub code	Notes	Example
6 Post feedback action *Actions that occur as a result of discussion*	Reification	Examiner adding notes to MS; examiner amending marks submitted for later (self) reference	EX: Okay, yes. I'm not doing this as I go along, I'm making notes.

Appendix E: Stop List [Pilot and Main Study]

- ; 1; 1a; 1ai; 1aii; 1aiii; 1b; 1c; 1di; 1dii; 1ei; 1eii; 1fi; 1fii; 1gi; 1gii; 2; 2a; 2ai; 2aii; 2b; 2c; 2d; 2ei; 2eii; 2eiii; 3; 3a; 3ai; 3aii; 3b; 3c; 3di; 3dii; 3diii; 3e; 4; 4a; 4ai; 4aii; 4b; 4bi; 4bii; 4biii; 4c; 5; 5a; 5ai; 5aii; 5b; 5bi; 5bii; 5c; 5di; 5dii; 6; 6a; 6ai; 6aii; 6b; 6bi; 6bii; 7; 7a; 7b; 7ci; 7cii; 7ciii; 7civ; 7d; 8; 8a; 8b; a; an; and; AO1; AO2; AO3; are; at; b; be; d; don; i; id; ii; iii; in; is; isn; it; iv; l1; l2; l3; ll; m; of; on; re; s; t; that; the; to; ve; y

Appendix F: Participant Consent Form

UNIVERSITY OF CAMBRIDGE
Faculty of Education

Team Leader Research Project Examiner Consent Form

One completed copy is to be retained by the participant and another will be retained securely by the researcher.

Full title of the project:	Team leader feedback research project
Name and position of the researcher:	Martin Johnson, PhD student, University of Cambridge

I confirm that :	Please initial the box
I have read and understand the information provided about the research *(MJ060214EIS)*	☐
I have had the opportunity to ask questions which have been answered fully	☐
I understand that my participation is voluntary and that I am free to withdraw at any time without giving reason, and if I do so I will inform the researcher	☐
I agree to allow my team leader to share with the researcher any feedback that they give to me through scoris, email, or over the telephone	☐
I agree to allow my team leader to audio record any feedback that they give to me over the telephone *(I understand that I can opt out of any recording prior to each telephone message)*	☐
In principle, I agree to take part in an interview. I understand there is no obligation to take part in any interview, and that I can withdraw my consent at any time	☐
I agree to the use of anonymised results and quotes in reports to the Cambridge Assessment Group and within the public domain	☐
I agree to take part in the study	☐

****

Examiner	*Date*	*Signature*
Martin Johnson	2nd May 2014	
Researcher	*Date*	*Signature*

Appendix G: Verbal Feedback Transcription Conventions [Main Study]

Observation	Annotation	Rationale
Pause	:::	A potential indicator of dispreferred response
Sigh	[Sigh]	A potential indicator of frustration
High rise	?	A potential indicator of surprise
Low fall	>	A potential indicator of disappointment
Exclamatory utterance	!	A potential indicator of emphasis or surprise
Lengthened syllables	Ye :: s	A potential indicator of emphasis or dispreferred response
Increased volume	A LOT OF	A potential indicator of emphasis or surprise
Stress	because	A potential indicator of emphasis
Overlap	A total of [40] ? [40] yeah	A potential indicator of agreement/aligned meaning
Contextual information	(checks mark scheme)	A potential indicator of concurrent features relevant to the discourse
Quote	When it says *start marking*	A potential indicator of concurrent features relevant to the discourse

Based on Jefferson (2004) and (Martin & Rouncefield (2003)

Appendix H: Discourse Codes [Main Study]

Superordinate code		Subordinate code	[System Name]	Explanation
Focus/ Interpersonal *Non-Feedback information*	[FI]	Opening	Opening	*Start of the interaction*
		Closing	Closing	*Closing of the interaction*
		Disagreement	focDisagree	*Conveying disagreement with marking; indicating a need to shift thinking*
		Agreement	focAgree/Verify	*Conveying agreement with marking; reinforcing appropriate thinking*
		Confronting gaps	focConfront/Gaps	*Dealing with crossed wires; different understandings*
		Distancing	intDistancing	*Moving 'self' from corporate interpretation* *Positive and Negative politeness strategies*
		Authority	intAuthority	*Allusion to power/senior examiner team*
		Social common ground	intSocial CG	*Social Common Ground: Reference to background knowledge of each other* *Information about 'the person' that is not specific to the task*
		Accentuation	Accentuation	*Highlighting issues (e.g. exclamation marks)*

Superordinate code		Subordinate code	[System Name]	Explanation
Giving information	[GI]	Directive information	giDirective	*Outlining what to do next*
		Marking principle information	giMarking principle	*Information on a general marking principle (e.g. 'a response would need to be... to get... ')* *Information on how marking needs to spread performances (psychometric focus)* *Modelling examiner practice* *Bridge forward into anticipated future experience*
		Privileged information	giPrivileged	*Examiner not present / closed information* *Insights into bigger overview which TL can access (e.g. 'lots of candidates do this...')*
		General information	giGeneral	*Non-specific/non-content related FB on particular marking (e.g. 'you're marking is good')* *Preparation/orientation for next information (e.g. 'the following comments...')*
		Technical information	giTechnical	*Convey system level information (e.g. how to deal with technical/procedural/administrative issues):* *'Technical learning' - how to use this system in this community* *Pointing out clerical errors (e.g. 'awarding a mark twice')/misplaced annotations*
		Mark statement	giMark statement	*Number only/statement of MS/statement of mark (absolute answer implied)* *(e.g. '[A] is [B]')* *Location of the mark is not clearly referenced*
		Location of credit	giLocation of credit	*Limited Perspective: pointing out that there is a difference rather than why... (No rationale included but location of mark is) Examiner needs to fill the missing information/concept to understand difference (e.g. 'We gave . for .') [rather than 'we gave for ... because...']*
		Rationale for credit	giRationale for credit	*Extended/General Perspective (Rationale Included and location is included)* *Making process of thinking/reasoning explicit - in connection with examiner error in context of a question (e.g. 'we thought that this meant...[general marking point]') - interpreting performance from TL view*
		Standards	giStandards	*Insight into standards/signs of 'quality'*
		Examiner rationale	giExaminer rationale	*Giving examiner point of view*
		Request information (Closed)	Request Info (Mono)	*Request for definitive information - monologic (expecting a particular answer [e.g. Y/N] - control is with respondent)*
		Request information (Open)	Request Info (Dialo)	*Request could be to establish dialogue/reciprocal transfer (e.g. 'What do you think?'/'Do you agree?')*

Superordinate code		Subordinate code	[System Name]	Explanation
Bridging	[Br]	Internal reference (Internal)	Bridge (internal)	*Reference to content within same message*
		External reference (External)	Bridge (external)	*Bringing together concepts across different documents e.g. Cross-reference to performance and/or MS wording (need to reference document to make sense of the message)*
		Historic reference (Historic)	Bridge (historic)	*Bringing together past and present information - contextualisation* *e.g. Cross-reference to previous communication/ Reference to previous/repeated marking issue*
		Cultural reference (Cultural)	Bridge (cultural)	*How is this thing 'seen' in another context? (E.g. how is this concept taught in school?)*
		Balancing perspectives	Balancing perspectives	*Recognising ambiguity/ complexity of decision making process* *Hints at the inability to capture elements of 'definite' outcomes in a MS/ across multiple examiners* *Underplaying 'difference'*
		Offer support	Offer support	*Making support role (rather than judgement role) evident* *(e.g. 'hope this is helpful?')*
Action	[Act]	Review	actReview	*Indication of review needed/carried out*
		Resolution	actResolution	*Indication of agreement with other's perspective* *Indication that interaction has reached a mutually understood conclusion*
		Reification	actReification	*Examiner adding notes to mark scheme*

Appendix I: Subcorpora Details [Main Study]

Description	Type	Size (words)	Corpus name
All feedback communication	Reference	95486	TL1-3E1-18_ALL_[Reference]
British National Corpus Academic	Reference	1032110	BNC aca
British National Corpus Spoken	Reference	1027807	BNC dem
All TL feedback communication	Subcorpora	84697	TL1-3_ALL
Examiner feedback communication only	Subcorpora	10789	E1-18_ALL
All email feedback communication	Subcorpora	81881	TL1-3E1-18_e
All telephone feedback communication	Subcorpora	13606	TL1-3E1-18_t
TL 1 feedback communication only	Subcorpora	15127	TL1_ALL
TL 2 feedback communication only	Subcorpora	23650	TL2_ALL
TL 3 feedback communication only	Subcorpora	45920	TL3_ALL
TL 1 and 2 feedback communication only	Subcorpora	38777	TL1 and TL2_ALL
TL 1 and 3 feedback communication only	Subcorpora	61047	TL1 and TL3_ALL
TL 2 and 3 feedback communication only	Subcorpora	69570	TL2 and TL3_ALL

Appendix J: Feedback Analysis Framework [Main Study]

Content analysis (line by line)	**L1; L2; L3 etc…**
Knowledge focus	K
Content information	*KC*
Technical information	*KT*
Perspective offered	Pe
State (definitive)	*PS*
Identify (locate)	*PI*
Describe (what seen)	*PDE*
Explain (rule)	*PE*
Clarify (build on past information)	*PC*
Direct (action)	*PDI*
Referencing Strategy	R
Internal bridge	*RI*
External bridge	*RE*
Historic bridge	*RH*
Cultural bridge	*RC*
Personal bridge	*RP*
Style	S
Emphasis U/D (Up/Downgrade)	*SEU/SED*
Contrast	*SEC*
Purpose (reference to global intention)	Pu

Appendix K: Non-Participant Semi-Structured Observation Schedule [Main Study]

Observed behaviour	Examiner	Message	

	Intentions: Projected common ground	Projected impact
TL	Is there a general or specific objective for this feedback? Is there an 'order of salience' to the information?	If you put yourself in the shoes of the examiner, if they saw a performance like this again do you think that they would be in a position to know how it should be marked as a result of this feedback message. What specifically would you like them to have learned, or developed, as a result of this feedback?
Examiner	Is there an element of the feedback message that you find most useful and why? Is there any element of the feedback message that you do not find useful and why?	If you saw a performance like this again do you think that they would be in a position to know how it should be marked as a result of this feedback message

Examiner: Message

Line1	Team Leader	Examiner
K		
Pe		
R		
S		
Pu		

Line 2	Team Leader	Examiner
K		
Pe		
R		
S		
Pu		

Line 3	Team Leader	Examiner
K		
Pe		
R		
S		
Pu		

Appendix L: Participant Information Sheet [Main Study]

Martin Johnson

c/o ARD

1 Hills Rd

Cambridge, CB1 2EU

Examiner Feedback Research Project Information Sheet *(MJ060214EIS)*

2/5/14

Dear *********

I'm undertaking a doctoral research project with the University of Cambridge that is funded by Cambridge Assessment (OCR's parent organisation). I am contacting you to let you know about the project and to ask your permission for help with some data collection.

The project explores the characteristics of remote feedback that a group of team leaders give to their examiners during examination marking. The project's aim is to support the development of good practice across team leaders at OCR in particular, and the Cambridge Assessment group more widely.

Your team leader on Unit F322 has agreed to participate in the project and I want to ensure that you are aware of how this might affect you. I would also like to gain your consent for some aspects of the data collection.

Feedback data gathering

I would like to collate the feedback given by your team leader to all of the examiners in their team during the summer 2014 examination session. This includes feedback given by email and over the telephone. This would allow me to carry out content analysis of these messages and gain a comprehensive picture of the team leader's feedback throughout the session.

- I would like your permission to allow your team leader to share the feedback messages that they give to you and any responses that you make to this feedback. *(I would access these messages through your team leader)*
- I would like your permission to allow your team leader to record any telephone feedback messages that they give to you. *(Your team leader would also offer you the opportunity to opt out prior to each recording)*

Examiner perception data gathering

I will also be carrying out an observation of your team leader whilst they are engaged in giving some feedback. Due to the responsive nature of marking feedback, it is not possible in advance of the observation session to know which examiners will receive feedback during this session.

- I would like your agreement, in principle, to take part in an interview if you do receive feedback as part of this observation session. The interview would allow you to talk through your reactions to the feedback. *(This interview would take place in a location convenient to you and last up to 2 hours. You would be paid around £44 to reimburse you for your participation time)*

You are free to decide whether or not to take part in the study. By participating in the project your anonymity will be maintained. Your anonymity will also be maintained in any reports that result from the project. The data collected in this project are stored in a secure, password protected computer system that only I can access. The study has been reviewed and approved by the Research Division of Cambridge Assessment and representatives of the Faculty of Education at the University of Cambridge.

I have enclosed two consent forms. If you are happy to participate in the project, please sign one and post back to me in the reply paid envelope, keeping the second form for your own records.

If you would prefer not to participate, please post back the forms without signing them so that I can update my own records.

If you have any further questions please do not hesitate to contact me.

Martin Johnson

University of Cambridge Faculty of Education

Tel: 01223553843

Email: mj415@cam.ac.uk

Appendix M: Feedback Quartiles and Keywords

Keywords for Quartiles 1 and 2[19]

n	Reference (173796 words)	Ref Corp %	Word	% Ch TLQ12	LL TLQ 12	% Ch TLQ34	LL TLQ 34		% Ch TLQ1	LL TLQ 1	% Ch TLQ2	LL TLQ 2	% Ch TLQ3	LL TLQ 3	% Ch TLQ4	LL TLQ 4
3	2132	1.23%	not	0.12%	**	-0.01%			0.20%	***	0.02%		0.19%	*	-0.19%	**
5	1802	1.04%	marks	0.17%	***	0.07%			0.25%	***	0.06%		0.35%	***	-0.19%	**
6	1754	1.01%	mark	0.12%	**	-0.04%			0.19%	***	0.04%		0.04%		-0.12%	*
11	1129	0.65%	here	0.16%	***	-0.03%			0.22%	***	0.08%		0.14%	*	-0.19%	***
14	809	0.47%	only	0.11%	***	-0.05%			0.18%	***	0.02%		0.03%		-0.12%	**
16	770	0.44%	no	0.06%	*	-0.01%			0.09%	**	0.01%		-0.03%		0.01%	
17	770	0.44%	we	0.09%	**	-0.02%			0.11%	**	0.06%		-0.04%		0.00%	
18	743	0.43%	script	0.05%	*	-0.07%	*		0.05%		0.05%		0.02%		-0.15%	***
19	739	0.43%	can	0.03%		-0.01%			0.00%		0.07%	*	-0.06%		0.04%	
25	660	0.38%	scripts	0.06%	*	-0.18%	***		0.07%	*	0.04%		-0.18%	***	-0.18%	***
31	589	0.34%	please	0.05%	*	0.01%			0.03%		0.09%	*	0.00%		0.01%	
35	517	0.30%	two	0.05%	*	-0.04%			0.09%	**	-0.01%		0.03%		-0.09%	**
38	502	0.29%	first	0.07%	**	-0.08%	**		0.05%		0.10%	**	-0.01%		-0.14%	***
42	453	0.26%	answer	0.05%	*	-0.02%			0.08%	**	0.02%		0.03%		-0.06%	
48	410	0.24%	some	0.03%		-0.03%			0.05%	*	-0.01%		-0.07%	*	0.01%	
54	390	0.22%	gave	0.06%	**	-0.03%			0.13%	***	-0.03%		0.03%		-0.07%	**
55	384	0.22%	second	0.07%	***	-0.10%	***		0.05%	*	0.10%	***	-0.07%	*	-0.13%	***
57	380	0.22%	when	0.02%		-0.03%			0.01%		0.03%		-0.07%	**	0.01%	
60	379	0.22%	or	0.00%		-0.03%			0.02%		-0.03%		-0.07%	**	0.01%	
66	331	0.19%	standardisation	0.08%	***	-0.12%	***		0.13%	***	0.02%		-0.10%	***	-0.13%	***
67	322	0.19%	correct	0.04%	*	-0.03%			0.06%	*	0.02%		0.02%		-0.08%	**
69	320	0.18%	them	0.01%		-0.04%	*		-0.02%		0.05%	*	-0.07%	**	-0.02%	
79	292	0.17%	scheme	0.00%		-0.03%			0.02%		-0.01%		-0.01%		-0.05%	*
91	268	0.15%	level	0.02%		0.00%			0.06%	**	-0.03%		0.01%		-0.02%	
92	267	0.15%	need	0.03%		-0.02%			0.04%		0.01%		-0.05%	*	0.00%	
100	248	0.14%	put	0.03%		-0.03%			0.03%		0.02%		-0.01%		-0.05%	*
103	237	0.14%	awarded	0.04%	**	-0.04%	*		0.09%	***	-0.02%		-0.03%		-0.05%	*
116	207	0.12%	explained	0.03%		-0.01%			0.02%		0.04%		0.03%		-0.05%	*
119	200	0.12%	next	0.02%		-0.05%	***		0.01%		0.03%		-0.08%	***	-0.02%	
120	198	0.11%	point	0.05%	***	-0.07%	***		0.05%	*	0.05%	**	-0.05%	**	-0.09%	***
122	194	0.11%	yes	0.00%		-0.04%	*		-0.01%		0.00%		-0.04%	*	-0.03%	

[19] An example of how the corpus analysis output can be interpreted to identify keywords is shown below. To the left of the focus word is a display of its frequency (number and percentage) in the reference corpus (i.e. all feedback texts). To the right of the focus word is a display of its frequency (percentage) in each comparative corpora, along with an indication of the significance of any difference between this frequency and that of the reference corpus. This element is crucial as it helps the analyst to determine the probability of whether the difference between the two corpora results happened by chance (Collins, Yao, & Borlongan, 2014). The log likelihood ratio indication (LL) is central to this interpretation. A log likelihood ratio greater than 3.84 indicates a confidence of 95 per cent that the result has not arisen by chance (i.e. a p value of <0.05) [indicated with * in this table]. A log likelihood ratio greater than 6.63 is a p value of <0.01, [indicated with ** in this table], and a ratio greater than 10.83 is a p value of <0.001 [indicated with *** in this table]. Finally, the per cent change value for each corpus indicates whether the word is overrepresented (a positive value) or underrepresented (a negative value) in the subcorpus compared to the reference corpus.

Keywords for Quartiles 3 and 4

Reference (173796 words)	Ref Corp %	Word	% Ch TLQ12	LL TLQ 12	% Ch TLQ34	LL TLQ 34		% Ch TLQ1	LL TLQ 1	% Ch TLQ2	LL TLQ 2	% Ch TLQ3	LL TLQ 3	% Ch TLQ4	LL TLQ 4
3087	1.78%	you	-0.02%		0.14%	*		0.01%		-0.07%		-0.12%		0.38%	***
1561	0.90%	have	-0.09%	*	0.04%			-0.17%	***	0.01%		-0.07%		0.15%	*
1330	0.77%	as	-0.01%		0.07%			-0.02%		0.00%		0.13%	*	0.01%	
1124	0.65%	was	-0.06%	*	0.16%	***		0.01%		-0.16%	***	0.24%	***	0.09%	
725	0.42%	just	-0.09%	***	0.08%	*		-0.13%	***	-0.03%		0.10%	*	0.06%	
709	0.41%	thanks	-0.15%	***	0.04%			-0.18%	***	-0.13%	***	0.05%		0.03%	
705	0.41%	if	-0.04%		0.05%			-0.07%	*	0.01%		-0.04%		0.14%	**
705	0.41%	your	-0.04%		0.07%	*		-0.02%		-0.06%		-0.01%		0.14%	**
695	0.40%	there	0.00%		0.07%	*		-0.01%		0.02%		0.10%	*	0.03%	
658	0.38%	they	-0.05%	*	0.10%	**		-0.09%	**	0.00%		0.04%		0.15%	***
646	0.37%	seed	-0.22%	***	0.47%	***		-0.33%	***	-0.08%	*	0.49%	***	0.44%	***
631	0.36%	do	-0.03%		0.04%			-0.06%		0.01%		-0.02%		0.09%	*
610	0.35%	which	0.00%		0.07%	*		-0.05%		0.05%		0.06%		0.07%	
565	0.33%	all	-0.05%	*	0.08%	**		-0.06%	*	-0.04%		-0.03%		0.18%	***
523	0.30%	hi	-0.11%	***	0.01%			-0.14%	***	-0.07%	**	0.05%		-0.02%	
479	0.28%	essay	-0.01%		0.08%	**		-0.01%		-0.01%		0.13%	***	0.03%	
460	0.26%	more	-0.03%		0.01%			-0.07%	**	0.02%		-0.02%		0.03%	
439	0.25%	very	-0.04%		0.04%			-0.06%	*	-0.01%		0.02%		0.05%	
432	0.25%	about	-0.05%	*	0.04%			-0.06%	*	-0.03%		0.04%		0.05%	
410	0.24%	marking	-0.04%		0.06%	*		-0.05%	*	-0.02%		-0.03%		0.13%	***
409	0.24%	also	-0.01%		0.06%	*		-0.03%		0.02%		0.08%	*	0.04%	
399	0.23%	has	-0.01%		0.02%			-0.01%		0.00%		0.08%	*	-0.03%	
397	0.23%	because	-0.01%		0.01%			-0.05%	*	0.04%		-0.02%		0.04%	
384	0.22%	well	-0.02%		0.07%	**		-0.01%		-0.03%		0.04%		0.10%	**
379	0.22%	look	0.00%		0.02%			0.00%		0.02%		-0.03%		0.07%	*
348	0.20%	best	-0.05%	**	0.06%	*		-0.07%	***	-0.02%		0.07%	*	0.05%	
348	0.20%	bit	-0.02%		0.06%	*		-0.02%		-0.02%		0.06%		0.06%	*
321	0.18%	think	-0.04%	**	0.03%			-0.09%	***	0.01%		-0.01%		0.07%	*
312	0.18%	know	-0.05%	**	0.03%			-0.04%		-0.06%	**	-0.05%	*	0.11%	***
295	0.17%	am	-0.06%	***	0.00%			-0.08%	***	-0.04%		-0.03%		0.02%	
284	0.16%	up	-0.03%	*	0.04%	*		-0.05%	**	0.00%		0.03%		0.06%	*
274	0.16%	regards	-0.04%	*	0.05%	**		-0.06%	***	0.00%		0.07%	**	0.04%	
272	0.16%	again	-0.04%	**	0.00%			-0.06%	**	-0.02%		-0.02%		0.01%	
258	0.15%	many	-0.04%	*	0.09%	***		-0.06%	***	0.00%		0.08%	**	0.10%	***
255	0.15%	enough	-0.01%		0.03%			0.00%		-0.03%		0.06%	*	0.00%	
249	0.14%	been	-0.04%	*	0.01%			-0.04%	*	-0.03%		-0.04%		0.05%	*
234	0.13%	had	-0.03%	*	0.01%			-0.04%	*	-0.02%		-0.02%		0.05%	
233	0.13%	last	-0.02%		0.02%			-0.05%	**	0.02%		0.01%		0.02%	
229	0.13%	really	-0.06%	***	0.08%	***		-0.09%	***	-0.03%		0.02%		0.13%	***
223	0.13%	through	-0.04%	**	0.00%			-0.04%	*	-0.04%		-0.05%	*	0.05%	*
221	0.13%	rest	0.00%		0.05%	**		0.00%		0.00%		0.11%	***	0.00%	
214	0.12%	tolerance	-0.01%		0.04%	*		0.03%		-0.05%	**	0.06%	*	0.03%	
214	0.12%	much	-0.03%	*	0.01%			-0.05%	**	0.00%		-0.04%		0.05%	*
201	0.12%	candidate	-0.03%	*	0.03%			-0.04%	**	-0.01%		0.07%	**	-0.01%	
195	0.11%	another	-0.02%		0.03%			-0.05%	**	0.00%		0.04%		0.01%	
190	0.11%	eg	0.01%		0.02%			-0.04%	*	0.07%	***	0.08%	***	-0.04%	
176	0.10%	etc	0.00%		0.02%			0.00%		0.01%		0.05%	*	0.00%	
175	0.10%	over	-0.02%	*	0.03%			-0.01%		-0.04%	**	0.02%		0.03%	
116	0.07%	thought	-0.02%		0.01%			-0.01%		-0.03%	*	0.00%		0.01%	

Appendix N: The Integrated Analytical Framework (IAF) and Building Common Ground

Features that support or undermine the building and maintenance of common ground. (**denotes elements that were not identified in the IAF)

Factor 1: Language Use

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Clear articulation	8 - To be awarded EE the answer just needs to have some evaluation, not necessarily three L4 ticks. (E24 M1 21.5.1806)	*To me it means that all the candidate must do is answer the question yes or no. So from that point of view that one word [just] was amazingly significant. (Y2 E24 SR interview 106)*			
Imprecise language use	The standardisation marking is pretty accurate. (E25 M1 26.5.1820)	*Yeah, I guess that word 'pretty'. Yeah, 'pretty accurate' is quite vague, and perhaps could have done with perhaps a slightly more specific comment. (Y2 E25 SR interview 32)*			
Cogent language use (to the point/ relevant)	7 - there is a comment mark for rivalry here as it is the reverse of the non-rivalry point in explanation and then a mark for saying it is a quasi-public good. (E24 M1 21.5.1806)	*[The TL's] explanation here is very short, but it's very, very accurately written I think. (Y2 E24 SR interview 88)*			

Specific referencing	The main problem is the essay Q7 Your are 9 marks too generous on this alone! (E16 M1 3.6.14)	*So panic, it's urgent it needs to be done. (Y1 E16 SR interview 42)*	Overemphasis**	Hope this is helpful. (E5 M1 24.6.14)	*[In my past experience, the use of exclamation marks implies that] you are being shouted at in an email. You can feel it. As a marker you get your feedback message and you can feel quite, particularly early on in the thing as you are building up confidence a bit, you feel quite a bit apprehensive. (Y1 E5 SR interview 37)*
			Incorrect referencing**	3b Very muddled answer, keeps referring to erosion, not given any credit, 0 marks not 3. 3aii Both points unexplained so only 2 marks given here not 6? 3b Long roots not allowed in cold environments because of frozen conditions etc, can accept large, widely spread etc though so only 2 marks not 4 Please put onto mark scheme (E18 M1 3.6.14)	*I'll open up 2443, 1a. Oh no, that's not okay, well maybe it's the other question, or it's not there, right, which one is it? And you know ultimately it's just a slip of the finger but it then means you probably spent maybe five minutes or so going through your other papers and by that point, you are just thinking 'which is the one?', and so, in that instance, that missing a number, probably, and you are already feeling a bit frustrating and stuff at that point. (Y1 E18 SR interview 104)*
The development of a stock of shared linguistic terminology	6 - no reference to what is driving this switch in	*It relates directly to something that's actually written in the mark scheme. (Y2 E24 SR*			

	production (ie. increase in demand) so we can't give the mark. (E24 M1 21.5.1806)	*interview 85)*			
Informal style encourages reciprocity and dialogue	Morning [examiner name], thank you for your first seed. (E1 25.6.1048)	*I think also it allows [the examiner] to feel involved. It's none of this 'Dear ... Yours sincerely', or whatever. [To] build up a rapport. (Y2 TL1 [E19E21E23] SR interview 350-364); Yeah, definitely. I do like that ... that sort of informality I guess, yeah. Yes, it's positive in a sense I think. Not so this is what you've done wrong. (Y2 E21 SR interview 39); I think [the way the team leader finishes messages with a ['thanks'] it's very efficient. I think thanks is just, it's something very small that doesn't really add much, but it just implies that 'yes, I am grateful that you're carrying on, I appreciate you're making my job easier', that kind of idea. So even though it's an automatic statement, it's still that automated sentiment as well which is nice. (Y2 E23 SR interview 43)*			

	This is out by quite a bit again but I can see the calls you have made (except for one mistake). (E8 6.9.1816)	*I use parenthesis rather than a comma to change the language, it has become more informal. It's like the brackets and the formality is sort of, I guess there is thought in it here, saying, "I feel like if I left that comment without the brackets and without dropping my tone a bit, it would sounds as if I am saying, "You are not giving this mark," whereas actually I relate to you, I can see why you didn't. (Y1 TL2 [E8E11E12] SR interview 106*			
	Hi [examiner name], I can give you feedback by Scoris messages. (E24 521.1653)	*[The use of 'Hi'] certainly feels like [the TL] is establishing a relationship with me that is more of a support and a friendly relationships rather than a critical relationship. (Y2 E24 SR interview 79)*			
Hedging reduces social pressure**	4a I think this loses the 4th sig figs mark and no minus sign. (E5 M4 27.6.14)	*I know bloody well it does. Rather than say 'It is' I say 'I think' because it is then inviting them. It's not confrontational. And it also opens it if they come back to me and say 'No, I disagree' then we can have a normal professional discussion,*			

		which is what [this examiner] would do which is fine. (Y1 TL1 [E2E5] SR interview 443)			
	Just received your 1st seed and I am afraid that it is once more far too generous so you are slipping back into old habits it seems. (E16 6.3.14)	*It's the human element isn't it? I just received this and it's not good you know [it's] bad news. "I am afraid". It's the human touch in a way isn't it and it's almost, "Oh I am sorry but can you just go back and check it?" You know it's fair, it's not accusing language, it's encouraging language, and it's kind of a teamwork language in a way. It's not saying "You are doing this, you are letting the team down". (Y2 E23 SR interview 6-10)*			

Factor 2: Content

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Elaborated reasoning focused on the task at hand	6a - pollution in the water ok for the example mark, and true cost of pollution not reflected by the private cost is enough for an analysis mark (just about getting idea of the price being wrong). (E8 M1 9.6.14)	*[The bit in parenthesis is there] because he is not going to get [that], that's not in the mark scheme, but here I am going priced at wrong level and I am saying, "It just about getting the idea of the prices at the right level." (Y1 TL2 [E8E11E12] SR interview 98)*			
		It's that patience in explaining the rationale that's so important to me, because now I think; well yes I can apply this. I get it. I understand your philosophy now. (y2 E24 SR interview 166)			
Lack of elaboration	1a(iii) we gave a mark here. (E5 M1 24.6.14)	*So that is probably the least helpful kind of feedback to get back. What [the TL] should maybe do is to go back to the script and then look at it from there and say 'So why have you given him the mark there?' The temptation is to ignore that a bit really and think 'Oh well you know…'*			

		because again you are on the timeframe and you just think it's probably one of those odd little things. Actually kind of crack on with the next lot. (Y1 E5 SR interview 57)			
Information allows participants to understand what is important to each other	6b - reduced overproduction gets this answer through Band 2 and it goes onto get 18. (E11 M1 8.6.14)	*Well it's just little reminders for me, again key phrases have come up where I have seen when I was practicing and going through the standardization process, certain phrases that helped me like you say kind of pull out from the mush, whether something hits a particular band, or notes saying some things are not acceptable. Again it's some words and phrases that I might have thought maybe they are adequate for a particular level and they are not so I've made a note of that. (Y1 E11 SR interview 77)*			
Language is meaningful (nuanced attuning)	7b No marks here - there is no mark for the IR absorption alone here - it is in 6c where this occurs and M can only score as an ecf when L is a ketodiol or an ald-diol, as per	*I've basically reiterated what's in the additional guidance, but in a slightly different way. (Y2 TL1 [E19E21E23] SR Interview 570)* *[It's the rephrasing that helps, that sort of iteration, that little movement around..?] Yes,*			

	Additional Guidance. (E19 M2 19.6.2047)	*yes, exactly, directly. Because it won't, it will not come again, I've understood it, let's do another way, another language, another way of saying the same thing. (Y2 E19 SR interview 251)*			

Factor 3: Time

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Reassuring and frequent communication (during learning)	1) Answers get capped at 10 marks if they don't get L3B2. L3B2, which should be denoted by a second L3 marker, can only be awarded if they explain how the subsidy corrects market failure (eg. solves the underconsumption problem). No L4 marks can be awarded without hitting L3B2. (E26 M1 21.5.1759)	*Yes, it is, it's reiterating. That says the same as that line. (Y2 TL2 [E24E25E26] SR interview 426)* *More frequent, good feedback, more frequent feedback. Through experience you've understood exactly what the mark scheme will accept or not accept so over time we don't need the feedback. (Y2 E19 SR interview 72)* *[Feedback is successful in terms of getting you more aligned] just with the sheer regularity of it. (Y2 E21 SR interview 117)*			
Involves less cognitive demand		*I think if it's not immediate they forget it when marking the scripts. (Y1 TL2 [E8E11E12] SR interview 62); If you have got a query about a paper you kind of want it resolved within 24 hours sort of thing really. (Y1 E5 SR interview 42); I would have marked this maybe the night before.*			

		Since then you had X, Y and Z to happen, and also you haven't just sent this one [script] but you sent four of them off. So [the TL] sent this back and I am thinking, "Right, I will go back to that script", and I am trying to work out what exactly [the TL] is talking about here, [the TL] is referring to erosion that kind of thing and I am thinking, "right which bit is that?", and it's just, it takes much longer, you have already got your own perception of what [the script] was worth so you have to then take perception away, literally delete what you have put on their previously (Y1 E18 SR interview 75)			
On-going and cumulative	I haven't seen any answer which actually makes the point about price falling without having first said supply increased (they're unlikely to have worked out price has fallen if they haven't worked this out!). So in the vast majority	*This is basically what I've said here, there's no point saying the same here again, so I've then had to give some sort of theoretical background and then gone back to really repeating what I've said before, hoping I've made more sense in the context. It's like well okay, you clearly don't see, understand my feedback on this specific script, so let's go a bit*	Language changes**		*I think in previous, previous years when we've used EE it's been to show an extended judgement and I think this time, well this time they're using EE to clarify that that is an evaluation. (Y2 E26 SR interview 50)*

	of cases for this mark yes there will be some explanation as to why prices have fallen. However, I think an answer which simply says "biofuel prices will fall which benefits consumers" entirely out of context could be awarded the mark. We didn't give the mark here because they didn't even link heating to biofuels - so overall we assessed the answer as not being relevant to the question set. (E24 M3 2.6.1922)	*broader and then let's come back. (Y2 TL2 [E24E25E26] SR interview 288)* *It's kind of a cumulative effect … which is why it makes sense. (Y2 E29 SR interview 97)*			
			Information overload	This is quite a long way out of tolerance, can you please review? (E11 M1 8.6.14)	*My default assumption is they are going to look at it because otherwise what's the point of feedback? In which case I don't need to waste [words]. The potency of the message goes in the more words you use in my opinion. (Y1 TL2 [E8E11E12] SR interview 75)*
Immediate	Examiner 5: The	*If you have got a query*	Time pressure	Examiner 29: For	*I don't need long blurb. If*

feedback**	only one I am a bit unclear on is the last one when the IR information mark is awarded. I thought they needed to specifically say that the broad peak is an OH linked to a carboxylic acid or is it sufficient to say there is C=O and an OH therefore must be a carboxylic acid? Morning [Name] Many thanks for your message re: seed feedback. I have re-read the script 650539832: - question 8b - and I think this answer has linked both the C=O and the OH to an COOH group and I would definitely award the mark here (E5 M3 26.6.14)	*about a paper you kind of want it resolved within 24 hours sort of thing really. I mean I was sending three, four kind of sets and I still hadn't got this feedback on the previous ones. Suddenly you then get a whole raft coming through. I think in that situation someone has got too much on their plate. (Y1 E5 SR interview 42)* *My view is, actually what they need is an immediate feedback, and if I frill it up and try and write too much, I will get through the feedback slower so one of the main, I guess intentions I am thinking of is I want this to be immediate feedback. (Y1 TL2 [E8E11E12] SR interview 62)*		question 4 a(ii) if the candidate writes salt pan for a(i) and then explains how its shape has been formed using water accurately, is it possible for them to get full marks for a(ii)? No because the landform is not found in this landscape. If we can we avoid double penalties but this time the PE decided not to accept these answers. (E29 M5 25.5 15)	*the answer can be yes or no then, you know … especially at this point. (Y2 E29 SR interview 53)*

Factor 4: Form

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Represented in a variety of ways including: A record of features of the interaction	There is one sentence in the first part of the conclusion where the candidate mentions the need for sustainable management but does not develop this and the rest of the conclusion is irrelevant. This mention gives him one mark only in AO2 – see mark scheme! (E16 M1 3.6.14)	*I find it easier when it comes in an email like this, because you can refer back to it. (Y1 E16 SR interview 47);* *I think this is really clear and in some ways it's quite nice to have it written down rather than just a phone call. I remember when I was doing it as a phone call; it was frantic scribbling at the same time. But there were points where it required me to think a bit more about them afterwards, so a discussion might have been better in some ways. [For example, in one situation] the reason I was making that mistake is because I missed out something in the mark scheme. I would have probably worked that out if I'd spoken to him about it. (Y2 E23 SR interview 6-10)*			
Represented in a variety of ways	4c Watch out for answers which talk about the	*So when they were talking about for the sensitivity, I was looking for that word*			

including: Social and visual cues	products being price elastic or products being insensitive to changes in the price, rather than supply. (E25 M1 26.5.1820)	*'proportional changes', but then I subsequently found out that, actually, they were OK with using the word 'sensitive' or 'insensitive'. Ultimately, the problem is that you've got the team leaders going to the standardisation with the principal examiner and spend a lot of time going through the mark-scheme. But it's those sorts of nuances of subtleties, of those discussions where you pick up on the odd comment here and there that aren't necessarily then formalised in the mark-scheme. [A couple of years ago] we did have a face-to-face meeting, and you do pick up on those little things. (Y2 E25 SR interview 77)*			

		Whereas the signals in a face to face meeting, I think you can do a lot more specific questions. You know, "Will this work?", "Will that work?", "Will this work?" Whereas it is very hard to do [that] by email. (Y1 E12 SR interview 22); It's those sorts of things where now, these words jump out at you, but when you're starting it's very … again, it's seeing the important parts. And again, if it was a conversation, then that would be emphasised automatically. (Y2 E29 SR interview 185); You feel in touch really. It's where you ask questions or if you see something that might get a mark or it gets a mark but you were unsure about it, you just sort of clarify and just there and then you can iron out the mark scheme. (Y2 E8 SR interview 21-23); It's the sort of thing that if you're face to face you can draw and show. (Y2 TL2 [E24E25E26] SR interview 219); What I was thinking is really in this, it would be much easier to do this on the			

		telephone or face to face like we are now because you could both look at it and we could discuss it. There is a slight problem with the time link of having to email and we haven't got the same paper in front of us. I mean, it is the same but it's not. (Y2 E29 SR interview 236)			
Spoken feedback is quicker to deliver		*There is no voice to this person, so it's just their comments and sometimes you feel a bit exasperated by them. On the phone, it's immediate, the person is very friendly. It's much better like that, and I actually think it probably took less time, for [the TL] to do that. (Y1 E18 SR*			

		interview 77)			
Questioning is part of richness**	I hope that this helps? (E18 M1 3.6.14	*So if he doesn't understand then I really would like him to get back to me. I think he would actually, if he is stuck. (Y1 TL3 [E16E18] SR interview 84)*			
Lack of 'rich' information	Essay 5 Put an insert mark at the beginning because no introduction for AO3. (E29 M3 29.5.1148)	*Sometimes I want to write something and I feel as I can't, I have to do it with an irrelevant or a question mark. (Y2 E29 SR interview 163)*			

Factor 5: Source

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Communication between trusted participants	…we were generous in a and gave the deltaH mark. (E2 M4 28.6.14)	*Because it's not a personalized mark scheme where I did it, it's the team. So it was 'we gave it' so 'we are wrong.' It's not a me against them, it's the team that is the hierarchy, if you wish to use that expression, gave that mark. (Y1 TL1 [E2E5] SR interview 222); I can see where you are coming from and if I was in the situation I probably would have given the mark myself and so I'm not going to be petty and get people's backs up. Some people do and I think it's totally unnecessary. (Y1 TL1 [E2E5] SR interview 240)*			
Shared information can lead to trust building		*And more recently there's things that have come up and [the TL] had said comments which have obviously shown that she has similar issues to me in terms of some of the things that have come up. And again, it's that I finally feel like having that relationship building at this stage that might*			

		have been helpful early on, but wasn't possible. (Y2 E29 SR interview 155); When [the TL] talks about asking [the Principal Examiner] for his advice, that's helpful because it makes you realise, it just puts a human at the other end of the computer. [Y2 E29 SR interview 155]			
Lack of rationalising links to distrust/ isolation	Generally really good marking but problems with the essays would mean you were out of tolerance on most scripts. (E26 M1 21.5.1759)	*I mean I've worked with other team leaders in other situations who have been really poor at giving feedback, I mean the worst thing is when someone rings you up and says, "Yeah, it's fine, it's fine", because you think well "No, no, it can't be" so in a way that makes me think "You haven't even looked at it". (Y2 E26 SR interview 128)*			
Mixed messages link to distrust	It's a messy answer but on reflection I agree we should have awarded an analysis mark for supply meeting demand. (E24 M1 21.5.1806)	*[The TL] told me on the telephone that when we've got that phrase, 'increasing supply meets increasing demand' then, because it's good as far as the Exam Board's concerned, you can award three marks because of the 'increasing supply' the 'increasing demand' and the word 'meets' that joins the two phrases*			

		together in the middle. But, they then give me a standardisation script and I mark according to what [the TL's] told me on the phone and they've ignored it. So that's why I found it very difficult then to apply some of the other principles that they were asking me to apply, because I thought, well how can I now that being the case? (Y2 E24 SR interview 55)			
Multiple sources	`5 - you can only award the luxury/necessity marks once - they are opposites so can't be presented as two different reasons, as it's the same factor. (E24 M2 29.5.1714)`	*Well, it wasn't actually the feedback that did that [changed my marking]. I had originally assumed that that was the right way to mark it and a colleague persuaded me that I was wrong. So I then marked it in the way in which my colleague marks it, and he's done a lot more of these than I have. And then [the TL] came back to me and said, 'Actually, you've got that wrong' and I thought, right I was right to start with, because it does say 'or' on the mark scheme and if I'd just followed my own gut instinct I'd have got it right the first time. (Y2 E24 SR interview 134)*			

Social pressure not to ask questions of authority – embarrassment	1ai I am not sure what you are doing here as you have 5 ticks for a start? Only give one tick per mark and look at the mark scheme to see how they are allocated. (E18 M1 3.6.14)	*You feel embarrassed, I suppose. You put five ticks because you are trying to make it perfect and you think, right she has told me before to only put ticks in the right place. I have done something wrong here, to the most simple question, that's the kind of thing that if one of my students did, I would call them up, and think they were a bit of a dumbo. (Y1 E18 SR interview 43)*			
Hierarchic relations can limit discourse		*Because I don't have a relationship with her and again it's not her fault at all, I felt that if I said, "look I have these circumstances this week", I'm worried about that so I didn't, and still haven't ever said that. Therefore you're both not working with the same understanding of what each other is expecting. (Y2 E29 SR interview 85)*			

Factor 6: Emotion

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Communication has a positive focus	In 2e(ii) we gave the second mark for 9.72/136.9 = 0.071 and in 5a we gave the first mark for the 'correct polymer with side links' (E2 M5 1.7.14)	*[I don't pass on all of the errors because] what you've got to avoid is them bringing other mark schemes into this, but you've also got to bear in mind their professionalism. (Y1 TL1 [E2E5] SR interview 78); his marking was still excellent so therefore that kept him saying 'Yes, I'm still a professional and I know what I'm doing.' (Y1 TL1 [E2E5] SR interview 93)*			
Negative content	If you could please review the comments I've made on each script, adjust as appropriate and submit that would be great. (E26 M1 21.5.1759)	*[The TL] has occasionally, not with every message over this session, but some of the feedback on standardisation scripts has been a bit more blunt. But then that can feel a bit of a jolt because, you know, things are going high, 'that's a great script', 'well done', 'that script is very accurate', and then you suddenly get a message that says "This one was out of tolerance, have a look at this again", maybe I'm just being*			

		sensitive but you kind of think "Oh no, I'm going wrong". (Y2 E26 SR interview 118)			
A professional face incorporates emotion		*It's quite a difficult thing being judged by your peers like this, because as somebody who has been teaching the subject for a long time, suddenly somebody is telling you that maybe what you view as being the way it is isn't maybe quite right. It's quite an emotional thing to receive that feedback. (Y1 E11 SR interview 43); It is slightly nerve wracking [to receive feedback]. (Y2 E29 SR interview 45); Well you're basically being told that you're doing it wrong or you've done that bit wrong, but as a "professional", in inverted commas, it's very difficult to take that. (Y2 E26 SR interview 122)*			
Positioning positive information	Thank you for your standardisation scripts which were excellent. (E23 M1 19.6.1045)	*In this case they were excellent, and so I told her so, rather at the beginning. So therefore, she wasn't on the back-foot, shall we say, when she was looking at it. (Y2 TL1 [E19E21E23] SR Interview 82)*			

		I've tried to build his confidence because I get the impression that he isn't a careless person and he's actually anxious. [The feedback]'s positive then, negative, rather than the other way round. (Y2 TL2 [E24E25E26] SR interview 66)			
Empathy**	Examiner 11: *OK. I didn't think it was clear enough for Band 2, but I'll ensure that I reward it if I see 'reduced.* Yes it's an easy route in really, but very few candidates are putting it so those who do should be rewarded. (E11 M2 8.6.14)	*He is not saying, "You are no good," he is saying I can kind of see why you might have not thought that overproduction is a good phrase to use, but if candidates are using that then yes, we think we should have it. By which point I have already agreed to do so, so everybody is friends. "It's an easy route," rather than "it is" feels more conversational. (Y1 E11 SR interview 110)*			

Factor 7: Recipient

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Encourages an examiner locus of learning control	I hope that this helps? I will just get the others quickly back to you and then I can let you be authorised for the 2nd batch of standardisation and these have to be accurate before I can finally authorise you for live marking. (E18 3.6.1055)	*"I hope this helps?", I know she has put a question mark there because she is trying to get some reflection across there and that actually does help to soften the blow a little (Y1 E18 SR interview 211)*			
	I agree this is a really marginal script and I'd be happy for it to be put through or capped. (E24 524.1830)	*I think the main point there for me was the fact that [the TL] was agreeing with me and I didn't feel like a complete idiot. (Y2 E24 SR interview 103)*			
		I haven't told her what the range is, she's going to have to go back and look at it and refer to the mark-scheme. (Y2 TL1 [E19E21E23] SR interview 546)			

View of professional examiners as controlling their learning		*[This examiner] is pro-education. It is not just a way of making money, this is an educational experience. There is a proactive dimension. (Y1 TL1 [E2E5] SR interview 35-40;* *Ultimately from my point of view I'm a teacher. I teach this course. I would like to think that people who mark it, the markers are professional and trying to do the best for the kids. (Y1 E12 SR interview 86);* *It's tempting to go, right got the salient points there, let's get on with this, let's do the next one because I have got 10 more to do and I just want to get on with it, and I did go back and look through them. I did, I did spend time, and I did spend time making sure I understood what [the TL] meant because for me, I am not doing it to try and make £4 a script to get that cash, I just want to make sure that I can teach it better next year. (Y1 E18 SR interview 214);* *[The examiner] is probably the most interesting because he's*			

		lovely, he asks a lot of questions and he's always messaging me. *(Y2 TL2 [E24E25E26] SR interview 24)*			
Social isolation**	The rest was fine. (E19 M4 24.6)	*There may have been one on there which was a bit iffy, but I wouldn't in this instance bring it up, because it could be a BOD [benefit of the doubt]. And so therefore, it's only going to make this particular examiner more insecure than she is already. [The message] does have an effect, because you're sat in a room like this, on your own. It's very impersonal.* *(Y2 TL1 [E19E21E23] SR Interview 740)*			
Lack of confidence**		*At the beginning [the TL] said "Oh you're obviously very experienced" which kind is quite a nice feeling for somebody to say, but I wondered in my mind "Oh does that mean that he's thinking I'm going to be fine", you know, kind of psychologically. Which was kind of a test at the beginning as an examiner because you're thinking*			

		"Can I do it? Can I get through standardisation? Am I actually any good at marking?" (Y2 E26 SR interview 111)			

Factor 8: Knowledge

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Elaborated reasoning articulates important codified and tacit elements of knowledge from the broader task context	We didn't give the mark here because they didn't even link heating to biofuels - so overall we assessed the answer as not being relevant to the question set. (E24 M3 2.6.1922)	*Here he's given me the rationale and for me, as I say, it's that patience in explaining the rationale that's so important to me, because now I think, well yes I can apply this. I get it. I understand your philosophy now and I know why the Board wants this, so that's what I'm going to do. It's very difficult to line up with something that you don't understand, isn't it? (Y2 E24 SR interview 166)*			
Feedback contains important context-specific information		*These bits here are the most important parts because, these paragraphs 2 and 3 [are important] for all the essays in the future (Y1 E18 SR interview 207)*			
Mark schemes contain implicit elements	We felt the judgement was developed as the idea of information provision being needed in the short term to stimulate demand with subsidies needed in the	*It's something that's not in the mark scheme and therefore you know, using that word "interesting" is deliberately vague because this is valid, it's right, but it gives a bit of space. (Y2 TL2 [E24E25E26] SR interview 162)*			

	long term to maintain that demand was valid and interesting. (E24 M1 21.5.1806)				
		If we had a face to face, we would have said, "Right look, when you see anyone using the technical words 'allocative efficiency and overproduction', unless there's something stupid with it, it's going to get band two." Whereas this is an easier concept, it's not technical so if you are going to do this you need to explain it properly but you can't write that in a mark scheme. (Y1 TL2 [E8E11E12] SR interview 206; The actual welfare loss reduction was not in the mark scheme. It hadn't been mentioned to me and so I didn't know. (Y1 E12 SR interview 57); [The examiner] isn't going to get that; it's not in the mark scheme. (Y1 TL2 [E8E11E12] SR interview 98)			

Factor 9: Feedback Giving

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Participants recognise and explore each other's perspective	6b - this was awarded 5 only, as it was felt there was no real application of how the tax works (it follows an incorrect statement that a tax reduces demand, then doesn't explain that tax shifts supply left, raises price, reduces quantity etc), so the answer was consideered to lack appropriate application. I can see the justification for awarding the mark you have though and think it could be justified in live marking.(E11 29.5.1057)	*[The TL] is basically saying it's an easy mistake to make and "Yes, I can see why you might think…" (Y1 E11 SR interview 105)*			
Based on ego		*[You said that you learnt about giving feedback from looking at it from the other side?] So, I felt like*			

		I didn't get given enough feedback, and so that left me feeling that by the end of the marking process I [was] not convinced that I understood the mark scheme fully. (Y1 TL2 [E8E11E12] SR interview 132)			
Involves decision-making		*We were generous on occasions because there are some that are a bit iffy, but then you go with the examiner. Absolutely, indeed. It depends on the individual. Again that is not a blanket statement because you have to know the people involved and their experience of marking. You do get to know them, you do get to know them. (Y1 TL1 [E2E5] SR interview 75)*			
Involves managing relationships	`Morning [examiner name] (E21 M1 23.6)`	*This is continuing a dialogue, so I've [just] spoken to him. I think also it allows them to feel involved. (Y2 TL1 [E19E21E23] SR Interview 350)*			

Factor 10: Pedagogy

Subcomponents that support common ground	Feedback extract	Interview narrative extract	Subcomponents that undermine common ground	Feedback extract	Interview narrative extract
Brings together varying perspectives	This is out by quite a bit again but I can see the calls you have made (except for one mistake). Can you please carefully review the below? (E8 M1 9.6.14)	*Rather than this sort of dressing up with more detail, I sort of, if I didn't understand the message I would ask, and that's what you are really supposed to do as examiners, Because, also remember, examiners have got the incentive to mark well, because they know there will be another seed coming next. So really my view is, communicate as regularly as possible as much information as possible, and if I've not given them enough detail they'll ask me. (Y1 TL2 [E8E11E12] SR interview 131)*	Transmission links to superficial learning/no link to examiner knowledge**	8a Must be a formula for water and the structure is interesting. (E5 M5 29.6.14)	*8a is quite interesting because in the mark scheme it does say for E and F allow H2O/water. And so I was giving a mark at that point if they put water. Now that is there in the mark scheme, however it would seem to be it must be a formula for water. So I start marking it to that point. Now in the previous year I probably would have gone back and queried that. But these days I don't bother because there is no point. If you are saying it needs to be the formula then I will just mark to that. (Y1 E5 SR interview 158)*
Indicates what experts look for	1e(ii) We are not penalising connectivity here. (E4 29.6.1957)	*'Do not penalize connectivity.' That's over here because that is important and shouldn't be hidden. (Y1 TL1 [E2E5] SR interview 202)*			
			Partiality allows construction (active learning)**	5b(iii) Not propan-2-ene. (E23 M1 19.6.1045)	*You can't give them too much, you do … would like them to go back into the script not just to*

					check. (Y2 TL1 [E19E21E23] SR Interview 128)

Appendix O: Keywords and Quartile Tables

Quartile 1: Locating and enforcing hierarchy

2-Left		1-Left			1-Right		2-Right	
					8	kept	9	this
				[we]	9	thought		
					6	gave	4	[credit]
					2	awarded		

We kept this in L2. [E8 26.5 1141]

We thought this was enough for the 1st mark. [E1 22.6 1010]

We gave credit for that. [E9 25.5 T]

2-Left		1-Left			1-Right		2-Right	
				[only]	2	gets		
					2	given		

Only given L1 not L2. [E14 13.6 0820]

First, note that the market failure diagram only gets L1. [E11 26.5 1206]

Quartile 2: Justification/rationalising

2-Left		1-Left			1-Right		2-Right	
		4	it		2	no		
2	as	6	there	[was]	7	just	2	about
		4	this		3	given		
8	the	4	answer		2	awarded		

5 - the second reason only gets two marks as there was no mention of agricultural products. [E27_602.1926_Q4]

We felt it was just about strong enough to get through L4. [E24_528.1028_Q2]

The justified judgement at the end means it was awarded 17. [E26_531.1815_Q3]

2-Left		1-Left			1-Right		2-Right	
		4	a	[marginal]	5	one	2	though

6b - a marginal one that we kept in L1. [E8 1.6.1651]

If they had said it eliminates over-production, or even that less production means less pollution meaning less harms from negative externalities, that would have been ok. A marginal one though. [E11 29.5 1057]

Quartile 3: justification/rationalising – bridging perspectives

2-Left		1-Left			1-Right		2-Right	
		7	I	[think]	4	it		
2	I	2	don't					

I don't think it gets to L3 until page 10 though, as page 9 isn't actually explaining a tax but just how market failure could hypothetically be solved. [E12 1.6 2032]

1a(iii) The 2nd mark is here - I think it is clear that the candidate is breaking the vdW's. [E1 19.6 0924]

2-Left		1-Left			1-Right		2-Right	
		2	they	[were]				
		3	you					

In 7d we gave the reagent mark as they were correctly given over the arrows of both incorrect equations. [E5 9.7 0954]

You were close with the marks on this question, usually just 1 mark over or under. [E15 27.5 0747]

6b - I'm not sure how you were awarding marks here. [E13 29.5 1148]

2-Left		1-Left			1-Right		2-Right	
2	can	2	see	[why]	4	you[v'e]		
2	this	2	is					

6 - we gave zero. I can see why you awarded three but we felt the idea that heating prices would fall was not well connected enough to the question in hand. [E24_602.1922_Q4]

4b No reference to particular species Agreed and looking at this again it does look as if this is why the mark of 5/6 was awarded, however it does not state that this is a prerequisite for L2 in the mark scheme. [E17 25.5 1820]

Quartile 4: locating, rationalising and mark reinforcement

2-Left		1-Left			1-Right		2-Right	
		2	suggesting		7	are[n't]	6	say[ing]
		2	whilst	[they]	6	don't	2	talking
3	even	4	if		4	sa[y/id]		
		2	though		3	have[n't]		

6b - doesn't get anything, because although VAT is right, it is put in the same sentence as income tax, suggesting they don't really have an understanding of what indirect taxes are. [E12 29.5 1111]

if they said overconsumption falls which reduces production which reduces harms from negative externalities such as pollution therefore reducing market failure that would be good enough. [E12 1.6 1601]

Only award if they are talking about what happens to demand when prices are lowered. [E13 3.6 1825]

2-Left		1-Left			1-Right		2-Right	
		5	ke[ep/pt]		29	[number/level]	2	marks
4	kept	3	it	[at]				
		2	capped					

3aii This answer was kept at 5 marks. Some loose explanation eg 'The debri in the glacier abrades and plucks the sides'? [E29_604.1228_Q4]

Yes, that's right - an essay that doesn't hit the L3B2 criteria will effectively be capped at 10. [E11 26.5 1135]

2-Left		1-Left			1-Right		2-Right	
2	enough	12	to		10	into	12	[band/level]
		4	doesn't		4	through		
		4	answer	[get*]	5	to		
		3	it					
		4	this					

6b - we thought produce less and therefore less negative externalities was just about TV to get into Band 2. [E11 29.5 1057[1]]

Yes they've made a silly mistake but clearly understand it. What follows about over-production means it gets through Band 2 as well! [E8 14.6 1419]

This means it gets the two explanation marks for rival and excludable, bringing the answer up to 4/6. [E24_524.1827_Q2]

2-Left		1-Left			1-Right		2-Right	
7	top	12	of		20	[number/level]		
4	bottom			[page]				
		4	on					

There is only one mark for both absorptions and the acid (on page 22) scores a mark. Hence, one less tick for the absorptions and one extra tick for the structure on page 22. [E3 18.6 1926]

Lots of good stuff should have been hi-lighted at the bottom of page 18. [E28_528.1451_Q1]

Does this get through band 2 on bottom of page 8, top of page 9? It is a strong answer but unsure whether it meets the requirements of band 2. [E12 8.6 1451]